McGRAW-HILL
PUBLICATIONS
IN PSYCHOLOGY

133

McGRAW-HILL PUBLICATIONS IN PSYCHOLOGY
CLIFFORD T. MORGAN, CONSULTING EDITOR

BEGINNING EXPERIMENTAL PSYCHOLOGY

McGraw-Hill Publications in Psychology

CLIFFORD T. MORGAN

CONSULTING EDITOR

Barker, Kounin, and Wright—CHILD BEHAVIOR AND DEVELOPMENT
Bartley—BEGINNING EXPERIMENTAL PSYCHOLOGY
Brown—PSYCHOLOGY AND THE SOCIAL ORDER
Brown—THE PSYCHODYNAMICS OF ABNORMAL BEHAVIOR
Cole—GENERAL PSYCHOLOGY
Crafts, Schneirla, Robinson, and Gilbert—RECENT EXPERIMENTS IN PSYCHOLOGY
Davis—PSYCHOLOGY OF LEARNING
Dorcus and Jones—HANDBOOK OF EMPLOYEE SELECTION
Dunlap—RELIGION: ITS FUNCTIONS IN HUMAN LIFE
Ghiselli and Brown—PERSONNEL AND INDUSTRIAL PSYCHOLOGY
Gray—PSYCHOLOGY IN HUMAN AFFAIRS
Guilford—FUNDAMENTAL STATISTICS IN PSYCHOLOGY AND EDUCATION
Guilford—PSYCHOMETRIC METHODS
Hurlock—ADOLESCENT DEVELOPMENT
Hurlock—CHILD DEVELOPMENT
Johnson—ESSENTIALS OF PSYCHOLOGY
Krech and Crutchfield—THEORY AND PROBLEMS OF SOCIAL PSYCHOLOGY
Lewin—A DYNAMIC THEORY OF PERSONALITY
Lewin—PRINCIPLES OF TOPOLOGICAL PSYCHOLOGY
McNemar and Merrill (Eds.)—STUDIES IN PERSONALITY
Maier—FRUSTRATION
Maier and Schneirla—PRINCIPLES OF ANIMAL PSYCHOLOGY
Moore—PSYCHOLOGY FOR BUSINESS AND INDUSTRY
Morgan—PHYSIOLOGICAL PSYCHOLOGY
Page—ABNORMAL PSYCHOLOGY
Pillsbury—AN ELEMENTARY PSYCHOLOGY OF THE ABNORMAL
Richards—MODERN CLINICAL PSYCHOLOGY
Seashore—PSYCHOLOGY OF MUSIC
Seward—SEX AND THE SOCIAL ORDER
Stagner—PSYCHOLOGY OF PERSONALITY
Terman and Miles—SEX AND PERSONALITY
Wallin—PERSONALITY MALADJUSTMENTS AND MENTAL HYGIENE

J. F. Dashiell was Consulting Editor of this series from its inception in 1931 until January 1, 1950.

Beginning Experimental Psychology

By

S. HOWARD BARTLEY

Professor of Psychology
Michigan State College

First Edition

NEW YORK TORONTO LONDON

McGRAW-HILL BOOK COMPANY, INC.

1950

BEGINNING EXPERIMENTAL PSYCHOLOGY

PREFACE

Experience in teaching experimental psychology has indicated the need for a textbook having the following characteristics: (1) short enough to be covered in one quarter or one semester; (2) dealing with principles by way of illustration rather than being a compendium of the findings of experimentation; (3) representing all the major areas in psychology rather than just one or two which happen to be the central interest of the author; and (4) designed so that assignments may follow the natural divisions of the material. For example, the best arrangement would be to divide the book into the same number of (small) chapters as there are class assignments in a semester.

Until very recently, there was no book on the market that could be called a beginner's textbook in experimental psychology, not to say one having the above-mentioned characteristics. Although the gap in the field of textbooks in experimental psychology is beginning to be closed, no text answers the description just given.

The present book was written to satisfy this demand. Many of the short chapters, each a single assignment in length, are occupied with an adequate description of one or two actual investigations such as are found in the scientific literature rather than with the usual cursory presentation of findings.

The text is not presumed to be an over-all coverage of each topic dealt with but rather a representation of each main division of psychology by one or more assignments, each consisting of an extended description of an exemplary investigation in that field.

In the production of this book I should have been helpless without the aid of Leola B. Bartley, who typed, retyped, and

examined the manuscript for inconsistencies and other errors. Grateful appreciation is due also to the various authors cited, and to the copyright owners, particularly the *American Journal of Psychology*, the American Psychological Association, Inc., and the Journal Press from which so much material was required.

EAST LANSING, MICH.
April, 1950

S. HOWARD BARTLEY

CONTENTS

PART V. PHYSIOLOGICAL PSYCHOLOGY

PART VI. COMPARATIVE PSYCHOLOGY

PART VII. LEARNING AND MEMORY

PART VIII. PREPAREDNESS AND ADEQUACY

PART IX. SOCIAL PSYCHOLOGY

PART X. LEGAL PSYCHOLOGY

PART XI. CHILD PSYCHOLOGY

PART XII. CLINICAL PSYCHOLOGY

PART XIII. INDUSTRIAL PSYCHOLOGY

PART I

HISTORY

CHAPTER 1

INTRODUCTION

This is a textbook in experimental psychology. It is written on the supposition that you can well begin to learn, without laboratory exercises, what experimental psychology is. Although laboratory work is indispensable at an early stage of your training, your first introduction to experimental psychology is well accomplished through verbal description and explanation provided by textbook and lectures.

The purpose of this introductory chapter is to explain what to expect in the succeeding chapters and how to regard the material therein. This chapter and the one following also provide a historical perspective of experimental psychology.

It is assumed that each chapter in the book is about the right length and degree of difficulty to constitute a single assignment. Whereas you are used to having the typical textbook tell you a little about much, the present text will tell you enough about a limited number of matters to make them useful examples. Most of the chapters will each be devoted to the description of but a single investigation such as has appeared in the journals of recent years.

How to Use the Text. You are to look upon the contents of the chapters that describe experimental work in light of the reasoning and procedures set forth in the early chapters on the methods and logic of experimentation. You should inquire into the investigator's purposes and how his investigation fits into the area which the work represents.

Some areas in psychology are much more difficult to penetrate experimentally than others. Some of them already contain much less dependable information than others. In some areas the

problem of control is much more difficult. In light of these factors, studies in these fields should be respected, even though they do not seem to be so tangible as those in others.

Among other things, you should be aware that there are two types of problems in psychology. The one type is the *population* problem, in which the question of adequate sampling is foremost. The other type is the study of mechanisms, in which a very few cases will generally serve to answer one's questions. Of late, population problems have become more prevalent. These pertain to kinds of people or ways in which people react. Hence they have much in common with out-and-out polls. For example, someone may want to know whether colorblindness (colorweakness) falls into a *specific number of types.* In contrast, one may wish to know about some *specific mechanism*, such as that of dark adaptation.

Certain questions are broad enough or general enough to be stated either as population problems or as mechanism problems. The one may be converted into the other by a change in the way the problem is stated. The prevalence in envisaging problems as population problems, and the frequency of dealing with questions where sampling is involved, have tended to induce some teachers of experimental psychology to stress highly complex statistical procedures.

The General Problem Confronting Experimental Psychology. The general problem which you confront in taking up experimental psychology is well illustrated by the following example. In the late sixteenth century, Galileo made a demonstration for the professors and students of the University of Pisa showing that objects of different weights fall with the same velocity. As you will remember, he dropped objects of different weights simultaneously from the Leaning Tower of Pisa. It is presumed that even before he made the demonstration, he had a good notion of what would happen. His colleagues, of course, all expected the objects to fall with different velocities. Why? Centuries previous, Aristotle had announced his law of falling bodies which stated that light objects would fall more slowly than heavy ones. This had set the mold of thinking throughout the centuries.

Let us examine the reason for disbelieving the dictum of

Aristotle prior to any demonstration. Assume, for example, a heavy body which we shall call A and a lighter one called B. According to Aristotle, A will fall faster than B. This can be subjected to a logical test. If A and B should be coupled together, as by a chain, they would be expected to fall faster than A by itself. But according to the law, B being lighter than A would retard A's fall. For this reason, A should not fall so fast as it would if falling unattached to B, or A, being heavier than B, would accelerate B's fall. It will be seen that these propositions are incompatible and therefore untenable. Accordingly, it must be concluded that all objects fall with equal velocity. This, of course, is just what Galileo's demonstration showed. Anyone could have determined this without the demonstration. If we were to assume that Galileo had a sufficient reason, even prior to the demonstration, to know what would happen, then we would have to say that he made his demonstration for "social" reasons —to convince his fellows when nothing else would suffice.

What does this situation illustrate for us? How does it apply to the study of experimental psychology at present? First of all, it reminds us that there are some propositions which, when examined closely, are meaningless. These propositions involve incompatible implications requiring opposing results. Propositions should be examined for this very purpose of seeing whether or not they have meaningless or absurd implications. It is obvious that if they do imply incompatibilities, they cannot be accepted as bases for experimentation.

Unfortunately the discovery of the absurdity of certain propositions sometimes tends to remain a private affair. This is to say, it is difficult to convince some persons of absurdities. To try to do so would be to collide with some of their deep-seated beliefs. Logical argument is often ineffective. Lip service to an idea is no guarantee that it has been accepted. Concrete demonstration of the "obvious" must sometimes be undertaken to compensate for the ineffectiveness of logical argument. Some "experimentation" must be undertaken for just this "social" reason. Thus, in general, reasons for experimentation vary all the way from this demonstrational demand to the need for discovering outcomes that can by no means be foretold.

This example, in which Aristotle was wrong, is not to be taken as an illustration of too great dependence upon reasoning. It is not that Aristotle thought too much; he did not go far enough or in the right direction with his thinking. In effect, Aristotle did not do enough *logical* experimenting. Had he gone through the operations of imagining the situations that we have mentioned, he would have seen that his propositions were incompatible and would have had to abandon his original law of falling bodies for a very different one. He would have made a discovery merely by performing the right armchair operations.

This example aptly applies to the present. The principles involved all the way through the example are as pertinent to the present as they were to Galileo's day. The example should alert you to begin thinking experimentally. Mere data collecting is far from experimentation and should never be thought of as the exclusive feature of research.

Perhaps this example should also help you to become sympathetic, at times, with certain "experiments" that seem to do no more than demonstrate the obvious. It is easy to scoff at them as being stupid. Before doing this, you should consider whether the individuals who performed the experiments had not become aware of the obvious, or whether they were convinced that a sizable number of colleagues needed the demonstrations.

Galileo's example also brings us to the mention of the usual clamor for *quantification.* There are times when a principle has been fully demonstrated by the results of certain operations, but the results do not seem to bear the weight they ought. The trouble seems to be that the results are thought little of because they are not quantitative. We have long been led to respect quantification almost regardless of whether what is quantified is relevant to the problem at hand. Galileo's demonstration involved no measurements in the strict sense. For his purpose, it would have been no more useful for him to have used elaborate and precise measuring devices than to have made the simple test that he did. It is to be admitted, on the other hand, that quantification would have provided data from which other kinds of deductions could have been made. An equation regarding the rate of fall as a function of time, etc., might thereby have been de-

termined. But the important thing for us to know is when and where quantification is required.

Early Experimental Psychology. When did experimental psychology start? How did it compare in its early days with what is called experimental psychology now? Psychology became experimental when it began to put its problems to some kind of test. Before it could do that, it had to approve of experimental methods and to see its problems as those which could be answered by them. Since this reliance upon experimentation did not become widespread overnight, the beginnings of experimental psychology are hard to fix. They cannot be pointed out in the same way that a historian can tell when the Magna Charta was signed or when Gandhi died.

To experiment with various aspects of human behavior came to be an accepted procedure less than one hundred years ago. Men, up to that time, were busy with those items in nature which they could know by simple and direct observation, could list, count, and classify like the number of toes on an animal's foot or the petals in a flower. Men, therefore, studied the human body long before they came to make planned experiments to answer questions about the way people feel and outwardly behave. Before they came to this, it had taken many centuries even to be willing and able to see accurately how the human body was constructed. For centuries men blundered past items easily detectable nowadays at a glance. Only by the seventeenth or the eighteenth century, for example, had a reasonably good understanding of the gross structure of the human eye been attained. When men did finally get to investigating human behavior, it was first by way of correlating sensory experience with body structure rather than by directly dealing with the kind of behavior problems which occupy our attention today.

During all this time there were certain individuals here and there who were curious enough about human behavior to speculate about it, and their speculations gave us what we had of our first psychology. They developed standard answers about the grosser features of overt behavior and about mind and concious-ness. This psychology was the creation of armchair philosophers. These rationalists exemplified the long-prevalent idea that man

could "believe" his way through all his problems. What was believed and announced by the leaders of one generation tended to be good enough, without modification, for the generations which followed. In this sort of thing, most of the time-honored institutions took a hand, and few of them have, even to this day, renounced authoritarianism for experimentalism.

The birth of experimental psychology waited until physiology had taken a definite interest in the nervous system and concentrated its attention on the senses. Finally, that day arrived. Following in the footsteps of the physiologists, certain philosophers took up the use of planned laboratory procedures for solving their questions. It could be said that experimental psychology came into existence at that time.

It is our purpose in this and the succeeding chapter to outline the highlights of what took place in experimental psychology from then on. Although it is a short history, it is packed with many kinds of endeavor and with an accumulated storehouse of detail. Much psychological experimentation has not been directed toward problems that the more socially-minded of today would call significant, for it has not concerned itself with personality conflicts, problems of motivation, anxiety, aggression, and other matters whose labels also float about in an easy fashion nowadays. Experimental psychology was more interested in how we feel the prick of a pin, with what part of the tongue we taste sugar, or how closely we can judge between two weights we lift. It was interested in sensation, attention, feelings, and modes and speeds of reaction.

Fechner. The man who has been given credit as the father of experimental psychology was Gustav Theodor Fechner (1801–1887). His life took in a large portion of the nineteenth century and was in part contemporary with a number of other famous names in science, such as Helmholtz.

There are several reasons for Fechner being called the father of experimental psychology. Circumstances had been leading up to the time when the technique of several of the scientific disciplines could be applied to the study of human behavior. It was in Fechner's day that various famous mathematicians were developing statistical theory. Included among these men were

Gauss, LaPlace, and Bernoulli. Herbart had conceived of the idea of threshold, or limen, and this was to be an important conceptual tool. E. H. Weber had performed some experiments on lifted weights from which he developed the idea that the just perceptible difference between two weights bore a fixed relation to the magnitude of the weight used as a standard. This ratio was later called the Weber fraction. For instance, if one weight is 1,000 grams and the weight that is just noticeably lighter is 980 grams, their difference is 20 grams. The ratio in question then would be 20/1,000 or 1/50. Weber found the ratio or fraction in lifted weights to be about 1/40 and in vision for length of lines to be between 1/50 and 1/100. In hearing, the ratio between two tones was 1/160. Weber believed that this ratio was independent of the magnitude of the stimuli used.

Now, since Fechner was a very methodical person and his interests lay in the direction of portraying the relation between physical events and experiential events,[1] the circumstances just mentioned went a long way toward giving him a chance to make the contribution that he did.

Fechner gave psychology much of its formal laboratory methodology. He distinguished between absolute and differential thresholds, or limens, and was interested in *averages*, extreme values, and variability about averages. On this account he applied the newborn statistics of his day to them. He also contributed the idea of *just noticeable differences*, or j.n.d.'s, the tool of *differential increment* by which he hoped to measure the magnitude of sensations.

In giving psychology a set of laboratory procedures (psychological methods and statistics), he gave psychology something very concrete to seize upon and use in the solution of its many problems.

Wundt. After the publication of Fechner's *Elemente* in 1860, the two next greatest events were the appearance of Wundt's *Physiologische Psychologie* in 1874 and the founding of the world's

[1] Fechner really was a pantheist and did not believe in a real distinction between conscious events and physical events; he used his experiments for the purpose of showing the fixity in the relation. He believed the mathematical equation for this would imply the unity in the two.

first laboratory of experimental psychology by Wundt at Leipzig in 1879. To implement this laboratory, he started a journal in 1881 called *Philosophische Studien*, which continued until 1903. To Wundt's laboratory came students from everywhere to learn of the new psychology and to do research. And it was in the *Philosophische Studien* that they published their results. The world's second psychological journal was the *American Journal of Psychology*. It was begun in 1887 by G. Stanley Hall. In this respect, America was not far behind Germany. By 1890 there was enough going on in Germany outside of the Leipzig laboratory to merit the founding of another journal. Ebbinghaus, with the assistance of König, started the *Zeitschrift für Psychologie*. This journal, too, was definitely devoted to experimental studies and had as its board of cooperating editors such men as Helmholtz, Aubert, Hering, Exner, Ewald, Preyer, von Kries, Nagel, and Zwaardemaker. These men, excluding Ebbinghaus, were physiologists. There were, however, a few psychologists who became associated in the enterprise. They were Lipps, G. E. Müller, and Stumpf.

Wundt's laboratory, during its first two decades, undertook the study of a number of kinds of problems, foremost of which were those in sensation and perception. About one-fourth of all the studies which were done during this period were in this area. Among them, studies on the various aspects of vision predominated. A number of investigations were made on the psychophysics of light and retinal excitation, several on the psychophysics of color, colorblindness, brightness contrast, apparent visual size, optical illusions, and peripheral vision.

Auditory sensation was the next most frequently studied area. In the analysis of beats, tonal intervals, tonal fusion, etc., came in for attention. In touch, problems of tactile localization and the 2-point threshold were prominent. The time-sense was studied through the estimation of temporal intervals.

Outside of sensation and perception, experiments on *reaction* (or reaction time) were undertaken most often. The work in this field seemed for a time to be a promising discovery, for it seemed to lay a foundation for the temporal measurement of the mind. It was assumed that *muscular reaction* pertains to the

perception of the stimulus, but not to its apperception. The *sensorial reaction* was thought to be similar to the muscular, except that apperception is involved. It was found that the sensorial reaction required more time to initiate; hence the time beyond that required for the muscular response must be the time taken for apperception. This technique involving the two kinds of reaction was called the *subtraction procedure* and was applied to still other further complicated reaction situations with the consequence that times consumed for cognition, discrimination, association, and will were determined. The field of *attention* was also included. Researches in this area involved such matters as fluctuation and range of attention.

Feeling was also given a place among the phenomena studied. During the nineties the laboratory was called upon to investigate the subject. Wundt had developed his new tridimensional theory of feelings, and it needed whatever support experimental findings could give it. It was during this time (1894) that the *method of paired comparisons* was developed from Fechner's method of impression. In addition to the method of *impression* in studying feeling, there was also the method of *expression*. It included the determination of pulse rate, breathing rate, etc., in relation to concomitant feelings.

Association was another subject of research at Leipzig. Almost as much work was done in this area as in feeling and attention. It, however, failed to find a permanent place in laboratory repertoires, for it failed to establish a definite form for investigation.

Students of the Leipzig Laboratory. Some of the men who became famous and whom we now know for what they did beyond their original work in the Leipzig laboratory are listed in Table 1, more or less in chronological order of their connection with it.

In addition to the Continental students who appeared at Wundt's laboratory, there were a number of Americans. Table 2 indicates some of them and the universities with which they became connected.

Experimental Psychology in Britain. There were several men in England who participated in the development of experimental psychology, even though at no time did the English universities

TABLE 1

Kraepelin	Munich	Abnormal psychology
Munsterberg	Harvard	Applied psychology
Lehmann	Copenhagen	Experimental psychology
Külpe	Munich	Experimental psychology
Kirschmann	Leipzig	Experimental psychology
Meumann	Hamburg	Experimental psychology
Marbe	Würzburg	Experimental psychology
Kiesow	Turin	Experimental psychology
Störring	Bonn	Psychopathology
Kreuger	Leipzig	Experimental psychology
Wirth	Leipzig	Experimental psychology

TABLE 2

G. S. Hall	Clark	L. Witmer	Pennsylvania
J. McK. Cattell	Columbia	H. C. Warren	Princeton
H. K. Wolfe	Nebraska	H. Gale	Minnesota
E. A. Pace	Catholic	G. T. W. Patrick	Iowa
E. W. Scripture	Yale, Vienna	G. M. Stratton	California
F. Angell	Stanford	J. H. Judd	Chicago
E. B. Titchener	Cornell	G. A. Tawney	Cincinnati

encourage psychology of this type. Because of this lack of encouragement, experimental psychology was rather late in getting a start there, and the roster is short. Although Galton, whom we must recognize as an early experimental psychologist in England, did his work contemporaneously with the Continent, the more extensive impetus for the development of experimental psychology possibly came from an expedition which Cambridge University sent to the Torres Strait in 1899. Included among the personnel were three men, W. H. R. Rivers, William McDougall, and C. S. Myers, for the purpose of making a psychological study of the primitive people to be found in the Strait. The duty of these men was largely to give mental tests, but they represented the reviving faith that anthropology must consider mental as well as anatomical characteristics. Rivers had previously done some work in visual perception, and on mental fatigue. McDougall was trained as a physician but took up physiological psychology after the expedition. In fact, the expedition expanded the interests of all three men.

The way was further paved for experimental psychology in Britain by Sir Charles Sherrington, the eminent physiologist who

in the nineties became interested in binocular brightness summation and in color vision.

In 1904, the *British Journal of Psychology* was started. During its first eight years it published one volume every two years. Of the nineteen articles published in the first volume, Boring, the present-day historian of psychology, counts seven as being experimental. Of these seven, five were by McDougall. During the same time a small amount of experimental psychology was being published in *Mind* and in the physiological journals. *Mind* was the first psychological journal ever published anywhere. It was started in 1876 by Bain, who represented the culmination of British associationism. But *Mind* was not primarily an experimental psychology journal, even at the turn of the century.

The first laboratory in England which could be said to pertain largely to psychology was Galton's anthropometric laboratory, which he set up in 1884. It was a demonstrational affair set up at the International Health Exhibition and later moved to the South Kensington Museum where it continued six years. To Galton (1822–1911) is to be credited a number of discoveries, techniques, and pieces of apparatus. He discovered synesthesia and number-forms. Among other devices, he invented a whistle now known by his name and used on man and other animals for determining the upper limits of audible pitch. He devised a pendulum arrangement for measuring reaction time and an instrument for measuring discrimination in depth of color. He also devised the Galton bar, an apparatus for measuring judgment of visual length or extent. During the short life of the laboratory, data on more than 9,300 persons were collected. These pertained to height, weight, breathing power, hearing, seeing, color sense, etc. In addition to his contributions to experimental psychology, he founded *individual* psychology (the psychology of individual differences) and originated the mental test. Galton was a very versatile man. Boring suggested that it was this very versatility which prevented him from being the father of a British school of experimental psychology. One does not now hear of his having a group of students clustered about him as Wundt did. Had there been such a group of learners, his influence might have blossomed out into a school of experimental psychology.

QUESTIONS

1. What is a population problem, and what is a mechanism problem?
2. What is *logical* experimenting in contrast to the more obvious kind?
3. What types of problems first engaged experimenters in psychology?
4. What developments helped to make the achievements of Fechner possible?
5. What were two of Fechner's specific contributions to experimental psychology?
6. Who founded the first formal laboratory of psychology in Europe?
7. What were some of its contributions in journals, ideas, and men?
8. Name two journals exclusive of the journal of the Leipzig group.
9. What was the name of the very first psychological journal published anywhere?
10. Who was the earliest contributor to British experimental psychology?

CHAPTER 2

HISTORY OF EXPERIMENTAL PSYCHOLOGY

Early Laboratories. At the end of the eighties and in the nineties a wave of founding psychological laboratories arose in Germany and in America. Although we do not know the exact time and number of them in Germany, the following can be said: By 1892, there were laboratories at Berlin under Ebbinghaus, Göttingen under G. E. Müller, Bonn under Lipps, and Freiburg under Münsterberg. At Munich, under Stumpf, and at Prague there were noticeable collections of apparatus on hand, but perhaps nothing which could be called laboratories. It was possible to do experimental work at Heidelberg, Halle, Strasbourg, and at Zurich.

In America, Boring tells us, there were at least 15 laboratories. James had developed a teaching (nonresearch) laboratory in about 1876 to drill his students in experimental psychology. At Johns Hopkins, G. Stanley Hall founded the first actual experimental laboratory. Although Hall had formidable aspirations for it, the laboratory soon lapsed for a time on account of Hall's shift to Clark University. By the beginning of the nineties, there were laboratories at Indiana, Wisconsin, Nebraska, Michigan, Iowa, Pennsylvania, and Columbia. In France the first laboratory of experimental psychology was founded at the Sorbonne in 1889 by Henri Beaunis and Alfred Binet.

The Eighties. During the eighties a variety of experimental investigations were made. The decade, the third in experimental psychology, marks the origin of various ideas which have since become well known. Of all the developments in this decade, eight will be mentioned. The first of these is von Kries' first book on the analysis of visual sensation in 1882. The second is the better known work of Stumpf, his *Tonpsychologie*, which

appeared in two volumes in 1883 and 1890. Not since Helmholtz's *Tonempfindungen* ("Sensations of Tone") in 1863 had there been anything of equal importance on hearing.

The third was the discovery of the spotlike nature of skin sensitivity when stimulated at or near threshold. M. Blix was led by his conception of Johannes Müller's theory of "specific energies of nerve" to explore the skin with styluses whose temperature could be varied. He wanted to locate specific skin receptors. In 1884 he published data regarding his discovery of cold and warm spots. Goldscheider made the same kind of a discovery and found also that pressure and pain manifest a pointlike distribution. He reported these findings in 1885. H. H. Donaldson, a student in G. Stanley Hall's Johns Hopkins laboratory, published, in 1885, his own findings on skin sensations, which confirmed those of the Germans. The findings of these men led to the view that there were four modalities of sensation in the skin rather than a larger and indeterminate number. This was a very important contribution.

It was in 1884 that William James formulated his original theory of emotion. He expanded it in 1890 and modified it in 1894. In the meantime, in 1885, a Dane named Lange published a quite similar theory. These men believed that the consciousness which is called emotional is nothing else but the awareness of bodily states which are induced by the perception of certain situations. While James was quite inclusive regarding the bodily states involved, Lange stressed vasomotor phenomena.

In France the study of feelings and concomitant bodily phenomena was taken up by C. Féré. His work led to the discovery of what we now know as the psychogalvanic response (formerly called "reflex"). Mosso in Italy also studied the expression of feeling. He is much better known nowadays for his ergographic studies of fatigue.

The sixth development was the beginning of what we now know as *animal psychology*, or a form of comparative psychology. Romanes was the leader in this through his urge, under Darwinism, to demonstrate mental continuity between man and animals. He developed the *anecdotal method*. During the eighties he wrote three books, *Animal Intelligence, Mental Evol-*

ution in Animals, and *Mental Evolution in Man*. Other publications in this field were made by Sir John Lubbock, J. H. Fabre, A. Forel, and Alfred Binet during this period.

The final two developments to be mentioned are the following: Preyer published, in 1882, the first book on child psychology which had ever been written; Galton formulated his theory of regression in 1886 and his theory of correlation in 1888. In so doing, he originated the kind of modern statistical methods used by psychology.

The Nineties. With the advent of the nineties the amount of work in experimental psychology expanded so much that it is difficult to describe briefly what was being done. There were several facts that can be taken as highlights. In these we follow what Boring has to say.

There was a decline of interest in the reaction experiment in its various forms. It was in this decade that animal psychology passed from the merely anecdotal to the experimental. It had originally been spurred on by the question of the validity of the theory of evolution. Now an interest in animal behavior (the animal mind) for its own sake supplanted the original one. A number of things were being done which we now see were a preparation for the mental-testing movement in America and France. There was also a great increase in the number of books written on psychology. Among them was James' *Principles of Psychology* (1890). A part of this general increase in the dissemination of psychological information was the founding of (1) the *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* (1890), generally known simply as the *Zeitschrift*; (2) the *Psychological Review* by Baldwin and Cattell in 1894; (3) *L'Année Psychologique* by Binet and Beaunis in 1895; and (4) the *Yale Studies*, 1892 to 1902, by E. W. Scripture.

Psychophysics passed from a subject of controversy and an object of improvement to being a set of tools, a method of research in solving the problems of the period. It was in this period that Cohn (1894) invented the *method of paired comparisons*.

Considerable attention was given to problems growing out of Ebbinghaus' original experimental work on memory. This was a period in which methods concerning the study of memorized

material were formulated. Ebbinghaus had earlier invented the *method of complete mastery*. The technique used in this method was improved upon by G. E. Müller and Schumann. Bergström invented the *method of interference* (1893). Münsterberg formulated the *method of reconstruction* in 1894. Baldwin and Shaw developed the *methods of recognition and reproduction*. Jost invented the *method of right associates*. Mary Calkins announced evidence for *primacy*, *recency*, *repetition*, and *vividness* as conditions of recall.

Although there was a reaction in some quarters against sensationism, the study of vision continued to prosper. Some of this study was impelled by interest in illusions, but a substantial amount was in the form illustrated by the work of von Kries. In 1894, von Kries formulated his Duplicity Theory of retinal function, a concept which today is extremely broad in import. G. E. Müller was another worker in the visual area and put forth at this time the concept of *cortical gray*. This was a theory to explain the fact that, even though supposedly there can be a balance between opposing retinal color processes, one does not see nothing, but sees gray—one has a positive experience.

In this decade notable research was done on cutaneous sensation. A revival of interest in problems of cutaneous localization and the 2-point threshold occurred. Max von Frey, in 1894–95, published his classical papers in which he emphasized the idea that tension within the skin, and not pressure upon it, was the essential feature in the stimulus situation. Von Frey established the existence of the phenomenon of paradoxical cold—the experience of coldness which is sometimes obtained with stimuli, the temperatures of which are higher than that of the skin. Victor Henri published a compendium on the work done in touch called *Über de Raumwahrnehmung des Tastsinnes* (1898) ("Concerning Space Perception from the Sense of Touch"). This was also an extensive analysis of the parts played by vision and kinesthesis as well as touch in space perception.

In the senses of taste and smell, notable work, if not a great deal of it, was also done in this decade. Oehrwall and, later, Kiesow worked on smell and established the four modes of taste that we now know—sweet, sour, salt, and bitter. Zwaardemaker

published *Die Physiologie des Geruchs* ("The Physiology of Smell") in 1895. Unfortunately, not much regarding smell has been learned since the days of Zwaardemaker.

In contrast to a certain amount of reaction against sensationism during the nineties, this period marked the time during which the number of estimates as to the number of sensations man could experience was most plentiful. For example, one estimate indicated the existence of about 11,000 tonal sensations. The estimates of visual sensations ran from a few thousand to several million and even to an indefinite number which might exceed this.

Titchener. Edward Bradford Titchener (1867–1927) was an Englishman who began his career at Oxford as a student of philosophy. Four of his five years there were spent in that field. The last year was spent in physiology under Bourdon-Sanderson. While there he translated into English the whole third edition of Wundt's *Physiologische Psychologie*. He did not publish this but took it with him when he went to Leipzig. There, Wundt told him that the fourth edition of his book was about to appear. Titchener set to work to translate it when it came out. But by the time that he had the fourth edition translated, Wundt was ready to publish the fifth edition.

When Titchener reached Leipzig, Külpe and the Americans, Pace, Scripture, and Frank Angell were there. Titchener's doctoral dissertation had to do with the binocular effects of monocular stimulation. When Titchener was ready to leave Leipzig, he found he could not return to Oxford, since little sympathy for experimental psychology existed there. The alternative was to come to America to Cornell University. So, in 1892, he came to the Cornell laboratory which Frank Angell had set up the year before. Angell was now leaving to go to the newly established Leland Stanford University.

The early part of his tenure at Cornell (which lasted the remainder of his life) Titchener spent in freeing himself from the philosophers and in translating the important books in German experimental psychology in order that there would be a basic literature in English for American psychology students.

Before the turn of the century, a number of students, who were later to become leaders in American psychology, had taken their

degrees with Titchener. For example, there were Washburn (1871–1939), Pillsbury (1872–), Bentley (1870–), and Gamble. In the same period, more than thirty studies from the laboratory had appeared. Quite early Titchener began his *Experimental Psychology*. The *Student's and Instructor's Qualitative Manuals* appeared in 1901, and the *Student's and Instructor's Quantitative Manuals* came out in 1905. Külpe acclaimed them as the most scholarly psychological production in the English language.

It has been said that Titchener never became a part of American psychology. But for a period during the latter part of his life and for a few years afterward, a sizable part of American psychology was influenced by Titchener through the work of his many important students. This was true despite the rise of behaviorism from about 1913 onward. Titchener took relatively little part in the affairs of the American Psychological Association which, to him, seemed a medium of the opposition group. Instead, he gathered about him his own group which, without formal organization, met annually from 1904 until the time of his death in 1927. This group was known as the Experimental Psychologists. In fact the group still meets but is now the Society of Experimental Psychologists, Inc., and is still limited in membership. Titchener was an editor of the *American Journal of Psychology*, and the work of the Cornell laboratory came out in this journal exclusively. The *Psychological Review*, founded by his contemporaries, was thought of as an organ of another group. This is reminiscent of the state of affairs in Germany with regard to Wundt and his journal and the group which sponsored the *Zeitschrift*.

Titchener did not confine himself to the study of sensation but later took up problems pertaining to attention, feeling, and the thought-processes. Accordingly, in 1908 his *Psychology of Feeling* and in 1909 his *Experimental Psychology of the Thought-processes* appeared. For him, feeling was an attribute of sensation and had but two qualities, pleasantness and unpleasantness. Stumpf had declared feeling to be one of the sensations, and Wundt had developed his tridimensional theory of feeling, attributing three pairs of qualities to feelings.

In 1909–10 Titchener published his *Text-book of Psychology*. While this book was too difficult for the college sophomores for whom it was written, it happened to be the only single book which contained his complete system of psychology. In 1915, he set out to revise the *Primer of Psychology*, which he had written in 1898. He renamed it *A Beginner's Psychology*.

One of the chief characteristics of Titchener's psychology was the application of pure introspective method for psychological analysis and the avoidance of what he called the *stimulus error*. He maintained that there was a distinct difference between observing sensory experience and judging object qualities. The psychological observer was to attend to sensations and not to objects as such. Sensations were to be described in terms of grayness, redness, loudness, etc. On the other hand, when observers included what Titchener called meaning, they were dealing with the stimulus object and committing the stimulus error. While this distinction was used to further his *sensationism* and in his search for sensory elements, it is a very proper distinction even today. We do not today, however, avoid problems in which the observer deals with objects, but it must be plain in every experiment whether or not the sensory effect on the observer, or the observer's identification of objects or object qualities, is to be required. For example, if a blindfolded subject is asked to respond when something touches him, he can tell what he thinks the object is or he can tell how it feels on the skin. In attempting to identify it, he is dealing with the "stimulus"; in describing his tactual experience, he is doing something very different.

Although some of Titchener's students are still active in psychology, they do not seem to be still pure Titchenerians. Their objects of interest are quite different than those in vogue twenty years or more ago. And the way they regard psychological problems has changed.

While Titchener and his students were busy with their form of experimentalism, there were others who did mainly one of two things. There were those who concentrated on psychological problems of a nature very different from those dealt with at Cornell. Cattell (1860–1944), who also had studied under

Wundt, had come back to America and become the progenitor of the psychology of individual differences. Raymond Dodge (1871–1942) made a very early study on reading, which gave valuable information on the range of attention and the nature of eye-movements. He spent a long life in this field and in that of the vestibular sense. Joseph Jastrow (1863–1944), who founded the experimental psychology at Wisconsin, did some critical work upon the psychophysical methods, anticipating others in substituting the *probable error* for the limen, or threshold, as a measure.

Then there were those who did considerable writing and thus influenced the course of psychology, but who did not go into the laboratory. The most noted of these was John Dewey (1859–) who was the founder of the school long known as functional psychology, and who, in 1886, published his *Psychology*, the first in America, and the second in the English language. The second of these was J. R. Angell (1867–1949), who fell in with functional psychology as opposed to Titchener's *structural* psychology. In 1904, Angell's *Psychology* came out. In the twenties, the psychology texts for beginning students included those of Angell, Titchener, Pillsbury, Seashore, James, and Griffith.

Experimental Comparative Psychology. The early work of Thorndike (1874–1949), which came out in 1898, is credited with starting the experimental movement in comparative psychology (the original name for animal psychology). Research laboratories for this form of psychology were established at Clark, Harvard, and Chicago, between the years of 1899 and 1903. Within the next decade there were at least eight such laboratories. Thorndike did not find evidence of inferential reasoning or insight in the behavior of his animals in puzzle-box situations. Lloyd Morgan called the type of learning displayed by Thorndike's animals trial-and-error learning. Thorndike also failed to find evidence of learning by *imitation*.

Just a couple of years later (1900), Robert M. Yerkes (1876–) began his studies of animal behavior and became the leader of the movement in animal psychology in America. All levels of animal life from lower animal forms to the anthropoid apes have come under his scrutiny.

John B. Watson (1878–   ) entered experimental animal psychology in about 1903. In 1907 appeared his work in which he reported on the white rat's ability to run the maze with various sense-modalities eliminated. He concluded that in running the maze, the rat is almost entirely dependent upon kinesthetic memory. Kinesthesis was the one sense he could neither control nor eliminate. The birth of behaviorism as a school of psychology dates from the publication of Watson's *Psychology from the Standpoint of a Behaviorist* (1913). In 1914, he published an *Introduction to Psychology,* and five years later his *Psychology* came out. It was a full textbook showing the scope of behavioristic psychology.

Among the others who figured largely in the movement of behaviorism were Hunter (1889–   ), Tolman (1886–   ), Dashiell (1888–   ), Weiss (1879–1931), and Lashley (1890–   ). Behaviorism has been pretty well absorbed into the main body of psychology, so that now it is not nearly so meaningful as formerly for one to say he is, or is not, a behaviorist. While it cannot be said that any of the above-mentioned men have forsaken behaviorism, the objects of their concern are not quite like they once were, and at least some of them show the influence of other movements which appeared in psychology then and subsequently. For example, Tolman's views may be said to combine behaviorism, Gestalt psychology, and purposive psychology.

Gestalt Psychology. Another form of experimental psychology which emerged in Germany and reached this country was Gestalt psychology, or configurationism. The founder of this school was Max Wertheimer (1880–1943), who took his degree with Külpe at Würzburg in 1904. In 1910 Wertheimer went to Frankfurt to Schumann's laboratory. Kohler and Koffka were there at this time and served as observers in Wertheimer's investigation of apparent visual movement. This was published in 1912, and its appearance dates the birth of Gestalt psychology.

The conventional view at the time of the birth of Gestalt psychology was that visual movement was a *series* of sensations which were synthesized and became a psychological process above the level of sensation.

Gestalt psychology was a revolt against the artificiality of such an analysis, an analysis stemming from sensationalistic elementarism. Wertheimer insisted that movement could be experienced immediately as movement. He looked upon it as being as immediate or as fundamental as any other experience. It was not to be accounted for by breaking it into parts. It could not be meaningfully reduced to a series of elements.

Other leaders of the configurational psychology school were Kurt Koffka (1886–1941) and Wolfgang Kohler (1887–). In 1921 this group, including Kurt Goldstein and H. Gruble, began the *Psychologische Forschung* as an organ of the school. In this country the school numbered among its adherents such men as R. H. Wheeler and Harry Helson. By now, however, the contributions of Gestalt theory to experimental psychology have become more or less absorbed within it. Gestalt psychology was never very well understood by most American psychologists, but in some sort of modified form its ideas, like those of behaviorism, have become the property of almost every psychologist.

Kurt Lewin (1890–1947), another one of the Gestalt group, progressed in a direction more or less his own. His efforts included the development of what is called topological psychology. This is a rational scheme which utilizes the logic of topology, a nonmathematical form of geometry. This sort of psychology deals with what Lewin calls man's life space and handles the relation of man to the field in which he is found. The group of concepts of which it is composed make it a very fertile device for dealing with man's social interrelationships. It is a promising kind of experimental psychology of today.

Experimental Psychology and Related Fields. During the late 1920's and the 1930's a number of those trained in experimental psychology did not seem inclined to follow the predominating path toward aptitude testing, personnel work, social psychology, clinical psychology, etc. They found opportunities more to their liking in related fields such as neurophysiology. A number of such workers occupied themselves with the study of the visual pathway and the auditory apparatus. Others turned to studying human behavior through examination of pre- and postoperative brain cases. Some of these workers were, for a

time, almost counted out of psychology, but it is becoming more obvious that they have contributed materially to general experimental psychology.

From its origin until the early thirties of this century, experimental psychology was occupied with quite concrete and narrowly conceived problems. Problems of personality, interaction between persons, or even the basic aspects of the individual's contact with his physical surrounds were not chosen for study. Isolation of interest was demanded, and on this account the findings of experimentation played little or no part in systematic formulations regarding personality and everyday human behavior. It was a period in which the more meticulous and exacting scientist could well say that personality was not his domain, for no one had the least notion as to what personality was. It may have been that the findings of the conventional studies of this period were meant to be building blocks out of which an understanding of human behavior was to evolve. But, if so, the investigators left the possible synthesis for others or for a later time. They were busy with accumulating repeatable (verifiable) facts. The last two decades of the 1800's and the early 1900's were a time of cataloguing and taxonomy. The experimentalists of this time were not aware of the ever-accumulating mass of problems concerning human conflict and tension. This demand was first met, as you now know, through the theories and practices of isolated groups in the field of medicine. Certain men in France, and later the Freudian school and its offshoots, occupied themselves with these emotional problems.

It took time for concepts from such origins to filter into academic psychology. At first the ideas, particularly of the Freudian type, were uncompromisingly resisted. But little by little the resisters took up with some of the ideas, if not in specific form, in a compromise fashion.

Social, abnormal, and then clinical psychology has become a predominant interest among psychology students. Interest in these areas along with a revolt against the sterility in the building-block sensationism has led to the study of man in his relation to the field in which he is found, and of the interactions between the field and the individual.

This study of man had its very crude, indefinite, and speculative beginnings, but now it is more and more apparent that these new interests of psychology call for the application of sound experimental method and the effort to envisage problems in solid scientific terms. It would appear that we are on the threshold of an era in which this is to be effectively accomplished. Already some of the more specific foundations of this method have been laid in the work of Kurt Lewin and his followers.

But there is yet another aspect of man's behavior that must come in for a rebirth in a new and more appropriate form. It is perception. A revival of interest in the study of perception appears to be on its way. This stems from demands of an increasing number of those in the remedial and healing arts (ophthalmology, optometry, reading clinics, etc.) who have come to see that a new understanding of vision is required. The study of vision (in schools training such practitioners) is passing from attention to it as a matter of optics to inquiry regarding the psychobiology of seeing. Psychology is being challenged in this area, and it remains for you who are now in the student stage to become aware of this challenge. It is you who will produce the answers to the urgent problems, if they are to be achieved in your generation.

QUESTIONS

1. Who founded the first experimental laboratory of psychology in America?
2. Who wrote the first book on child psychology?
3. Name a very early worker in animal psychology and tell what his objectives were.
4. What is the best known theoretical contribution of William James?
5. What did Ebbinghaus contribute to experimental psychology?
6. What did Zwaardemaker contribute to sense psychology?
7. Who did the early notable work on the cutaneous senses?
8. Name some contributions of Titchener to experimental psychology.
9. Name three of Titchener's students.
10. Who were the early men in behaviorism?

PART II

NATURE OF EXPERIMENTATION

CHAPTER 3

THE BASIS OF EXPERIMENTATION

In the foregoing chapters sketching the history of experimental psychology, it was pointed out that it was *experimental* psychology which was expected to take a place among the other biological sciences. It still is only the employment of scientific method that provides us kinship with other sciences. Students of psychology should, therefore, interest themselves in psychology as a science and not as a clutter of curiosities or as a panacea for their personal troubles. The "self-help" aim should be no more emphasized today in psychology than it is in the study of chemistry or botany.

Whereas the distinction between experimental and other kinds of psychology in terms of "field" became conventional for a time, the beginning student of psychology is now to think of experimental psychology as any and all psychology occupied with soluble problems and employing true scientific method in solving them.

In the present chapter, the essential nature of scientific experimentation is to be made clear. In anticipation of this, it may be said here that experimentation is as much a way of thinking as it is an outward set of laboratory procedures and cautions. Generally, experimental thinking involves a set of operations that are open to the eye, but there is nothing about them which distinguishes them from other outward procedures. To arrive at their significance, one must examine the thinking that determines them. This alone will provide the correct idea of what the procedures mean.

To point out the earmarks of a scientific experiment is the aim of the present chapter, hence the student is to be on the lookout for these.

Experimentation. The deliberate accumulation of knowledge depends upon experimentation. Scientific method is the method of performing experiments. On this account it is needful for us to consider what the essentials of experimentation are. We must know exactly how to distinguish a scientific experiment from all other kinds of procedures.

An experiment has often been said to be a set of conditions and procedures which will yield repeatable results. Granted that randomness must be minimized in experimentation, this statement misses the central point.

The concept of what experimentation is and what constitutes an experiment varies considerably from person to person. A great diversity of procedures has been given the label of *experimenting*. In the simplest form it is thought by the common man to be merely that of *trying* or *trying out*, or doing something the outcome of which is dubious, unforeseeable, and not predictable. For him, experimenting contrasts with all forms of routine activity leading to known ends which have been attained before. A concept of experimenting so general as this will, of course, not do in science. A more formal set of requirements must be looked for.

Bentley, in the Golden Jubilee Volume of the *American Journal of Psychology* (1937), has given an analysis and criticism of what experimentation has come to mean to various men in psychology.

Among the many meanings now attached to the term "experiment," he found he could distinguish at least eight. They are as follows: (1) to try something (this is the general concept already mentioned); (2) to make inquiries and seek opinions; (3) to perform certain gross operations and to translate them into the conceptual jargon of the day; (4) to accept a prejudice or propose a theory and then to find examples to illustrate its operation; (5) to watch a natural event and guess at its causes; (6) to plan and arrange an event and the conditions for it; (7) to solve a scientific task by the procedure just stated, and (8) to submit the results of this solution to a critical interpretation in the light of sufficient and precise knowledge.

It should be pointed out that these paradigms represent not only various *kinds* of procedures but also various degrees of com-

pleteness of performance relative to a possible problem situation. For example, procedures 7 and 8 might well be taken together to constitute a more complete procedure.

Bentley calls procedure 1 the *haphazard* or *street* method of "experimenting." Any course of action which is entered into without planning partakes of this characteristic. Bentley provided a very apt illustration of this way of doing in the behavior of the dyspeptic who "experiments" with his own health by omitting a meal per day, in taking an added amount of exercise, or in taking a cup of hot water each morning before breakfast.

The second procedure sometimes called "experimenting" (in psychology) is exemplified by the classroom teacher who passes out sheets of paper asking for answers to one or more questions, such as attitudes regarding war, food preferences, preferred colors of neckties, etc. He may also under these "experimental conditions" elicit attitudes and prevailing prejudices upon matters of taste, morals, or politics. To complete this form of "experimenting," it is generally demanded that certain statistical derivations (averages, mean deviations, etc.) be obtained, or even that factor analysis be performed. Whereupon the numerical products are taken as "demonstration" of the "scientific nature" of the procedure.

The third performance called experimenting is exemplified in requiring certain reactions from members of a group, whereupon they are quantified or classified and said to be evidences of certain conventional abilities or traits, such as "immediate memory," or "cooperativeness," etc.

The fourth performance is, of course, the actual assuming of the existence of some process or set of relations and then setting out to find instances of it. The main criticism of this procedure pertains to cases in which a few instances are taken to establish the validity of the theory. We may call this proof by selective demonstration.

The fifth procedure, the guessing at the cause of a natural event, is illustrated by the method of free association, where the experimenter declares that cerebral blocking or an inferiority complex accounts for the subject's behavior.

According to Bentley, experimental procedure should involve the following: (1) a task clearly and significantly formulated; (2) an adequate knowledge of the status of the problem and pertinent results previously obtained; (3) the general and particular implications of the problem in a broad scientific context; (4) a command of the situation to be set up, which involves thorough training and experience; (5) an expert appreciation of sources of error and ways for doing away with them; and (6) the ability to carry out a critical evaluation of the data obtained.

Experimentation starts with a problem. Genuine experiments are *planned*. Every experimental plan evolves from two kinds of assumptions. They are the experimenter's doubts on the one hand and his beliefs on the other. The assumptions which he makes but which he doubts are called *hypotheses*. The assumptions which he makes that are believed to be true are called *presuppositions*. The latter are premises upon which the experimenter proceeds.

The experimenter chooses the hypotheses he makes because they are in keeping with the facts of his own observation and those of others about which he happens to know. Since both his own and other people's observations are limited, the hypotheses may collide with facts that will later be disclosed. The experimenter makes his hypotheses also because by their use as premises he can deduce certain items capable of being verified or found to be false. In a sense, the experimenter logically projects ahead of his information at the time. His procedures from that point have to do with testing his hypotheses. Their substantiation, however, does not consist in seeking those facts which seem to be in line with them, but rather those which seem actually to contradict them. He should search for those facts which, if they can be found, will disagree with the hypotheses and lead to their abandonment.

The hypotheses the experimenter makes generally cluster into a single set of interrelated assumptions and on this account are often spoken of in the singular. An early step for the experimenter is to examine the hypothesis he has made to make explicit its logical implications. Naturally any hypothesis, to be materially true, cannot imply any material errors. Although a hypoth-

esis may consist in a set of ideas perfectly logically interrelated, this self-consistency does not determine their material truth.

After making explicit some important implicates, the experimenter is ready to hunt for facts which, if they exist, will be contrary to the implicates and thus make them untenable.

More subtle than the hypothesis the experimenter forms are his presuppositions. It will be recalled that presuppositions have been defined as the assumptions which are *not* doubted.[1] It is obvious that there are many more conceivable assumptions meriting doubt in one way or another than those which may be safely taken for granted without the least question. The experimenter most likely verbalizes his hypotheses but seldom takes equal trouble with his presuppositions.

A very frequent reason for the experimenter not to doubt the presuppositions he makes is that he is not aware of making them. Presuppositions are generally merely implicit components of the set of hypotheses the experimenter forms. In such cases, they are called *suppressed premises*. In any case, the experimenter reasons from them, and it is most important that he happen not to doubt them, and that the experimental procedures, whatever they may be, are planned in such ways as not to test them. As far as he is concerned, they are perfectly free of suspicion.

The two kinds of assumptions just described form the framework for the actual procedures which ordinarily are spoken of as experimentation itself. From what has just been said about this logical framework, it can be seen that the experimenter's thinking is a crucial ingredient in an experiment. Its role must be fully understood. It should be obvious that the assumptions do actually determine (1) whether a given experiment is to be undertaken; (2) what course the investigation takes; (3) and what the interpretation of the data will be when once gathered.

The thinking involved in scientific experimentation must always pass the test of logical consistency and be premised upon

[1] Those presuppositions which are not open to proof are called *axioms*. Axioms are necessary in building any kind of a system of thought and are at least remotely implied in experimentation. People disagree on axioms, but just so long as it is clear what the important axioms are, systems can be inspected and understood.

material fact. When the thinking involved is unassailable, the overt procedures will most likely be correct. Correct thinking makes due allowance for the possibility of not possessing certain information. Complete knowledge is not required before experimentation can be planned and executed.

The assumptional constituents of an experiment can be illustrated by the following example. The experiment is typical of very many that have been performed in the past and are still being performed. Let us say that an experimenter wants to know the recuperative effect of 8 hr. sleep per night or to demonstrate the existence of impairment which he supposes occurs in everyday activity. It hardly matters which, for the experiment will likely be planned in about the same way in each case. As is commonly done, he will probably devise some sort of a standard task for his subjects to perform at a fixed time shortly before going to bed at night and shortly after getting up in the morning. He will compare the performance made at these two times of day. The experiment will be continued over a protracted period so as to make it as representative as necessary. A difference in some quantitative aspect of the performance will indicate to the experimenter the impairment incidental to daily activity and the recuperation brought about by sleep. Whether the underlying reasoning is valid or not will be seen as we inspect the presuppositions and the findings of certain experiments conducted in the manner just described.

The presuppositions involved in the experiment are as follows:

1. That individuals become impaired by the usual everyday waking activities.
2. That impairment will affect some performance chosen as a test.
3. That this effect will be detectable as a change in amount of accomplishment in the test-performance.
4. That sleep reduces impairment, which reduction is also detectable.
5. That the restorative effect of sleep can be detected shortly after waking.

It is possible to extend the number of presuppositions, but in so doing, we run the risk of including implicates or corollaries

of those already listed. The presuppositions already stated will serve very well to show the principles involved in the thinking underlying an experiment such as is in question.

We need next to know what kind of results have been obtained from experiments of this kind. Both the experiment just described and others very similar to it have already been performed. It was found in several of the experiments, much to the experimenter's surprise, that the subjects performed definitely better at night, shortly before going to bed, than they did in the morning, shortly after getting up. Using the presuppositions listed above, we should have to conclude that *the greater the impairment, the better the performance.* Most people are not willing to accept this. They do not believe that in order to get individuals to perform best they should be impaired. Most experimenters are not willing to accept the conclusion just stated, even when they are among those who had accepted the original presuppositions. It is obvious that those who did not accept the original presuppositions need not accept the above-stated conclusion.

What must the individuals do who made the presuppositions? Naturally, they must make some change in the presuppositions, or else accept the conclusions.[1] In discarding the conclusion, what change or changes must be made in the presuppositions? This is a tricky matter, for the assumptions now in question were taken as the premises of the procedure whose outcome is now doubted. What ground does the experimenter now have upon which to stand?

It actually turns out that if he doesn't accept the conclusion, "The greater the impairment, the better the performance," it means he was all the time presupposing that "the greater the impairment, the poorer the performance." This means simply

[1] Let it be known that in spite of the unwillingness to accept the conclusion that impairment improves performance, few, if any, of the conventional experimenters upon impairment have openly and explicitly remodeled their presuppositions. Any changes that have occurred can only be deduced in the more recent behavior of these individuals. Such is the subtlety of the reasoning that goes on in human thinking. It is for this reason that cautions about reasoning are so important and are occupying our attention here.

that his presuppositions did not form a group which would not collide with material fact. It is hard to tell just how far back in his thinking it would be necessary to go in order to arrive at internal consistency and at assumptions he could rest upon. It should be at least evident from what we have said that the planning of a worth-while experiment is no trifling matter.

It is obvious that, to begin with, the experimenter could have held presuppositions quite different from those we have listed. For example, it might be assumed that impairment, except under unusual conditions, would be too slight to affect a test-performance one way or another. Or, it might have been taken for granted that, even though considerable impairment occurs, it could not be *dependably* detected by reduced accomplishment in a test-performance. It might have been assumed that impairment would be detected only by methods suitable for detecting cellular behavior. The clash between such presuppositions, of course, involves *definitions of terms* such as "impairment." But the differences in the presuppositions actually involve more than mere definitions, for we could as well substitute Greek letters or other symbols for terms such as "impairment" and still find that the differences exist.

Numerous other experiments have been performed under presuppositions such as we have listed, in which the effects of alcohol, narcotics, exertion, deprivation of oxygen, etc., have been observed. In each case impairment, considered as some kind of tissue effect, was imputed to underlie outward changes in performance. In each case, the measure of the impairment was sought, under the supposition that when the impairing agent is in operation, its effects will become manifest in some *decrement* of performance. It was also supposed that the *degree* of decrement is related directly to the degree of impairment. Impairment was supposedly related in simple fashion to the concentration of the agent in the blood. Actually, little correlation between the three items has been found, and in light of this, interpretations have varied widely.

In virtually all studies made on impairment, the possibility of *compensation* has been overlooked. This is equivalent to saying that the experimenters have proceeded under the assumption

(suppressed premise) that compensation does not occur. It has been shown that compensation is one of the very common characteristics of human activity. By compensation is meant a change in the manner of performing a task. This may include shifts from the use of one set of muscles, or one mechanism, to others in accomplishing the same end. Compensation is the changing of the means to an end. Activity goes on but in a different way. The measured results do not reflect the shift in means. Had the idea of compensation been included in the set of presuppositions in planning the experiment described above, the experiment could have been fashioned very differently, and the results therefore would have been quite different than reported.

The Solution of Problems. It is to be kept in mind that experimentation performs the function of answering questions or solving problems.

It is possible to ask two kinds of questions: (1) those which by their very nature cannot be answered by experimentation and the processes of logic, and (2) those which can. If one must have an answer to a problem in the first class, he must seek it by processes that satisfy his feelings and his beliefs. What will satisfy one person in such a case is not likely to satisfy someone else. The satisfying answer, whatever it may be, will evolve and take shape in keeping with the questioner's already-constructed system of beliefs. It is obvious that we are not interested here in problems of this sort.

It is sometimes possible to transform a question of the first class into one of the second. This is particularly true when the question belongs to the first class simply because its faulty construction puts it there. Such a question is not quite like those which arise within a system of faith. It is simply an improperly expressed question, and if taken literally, its answer can arise only out of the nonreasoned type of faith.

Three ways of classifying problems of the second class are: (1) with reference to their worth, as some questions pertain to little that is important to anyone or to anything from which generalizations may be made; (2) with reference to whether they are "neat" (easily solvable on account of involving laboratory

conditions that can be well controlled), or more complex and troublesome, but very important to society; and (3) with reference to their breadth. In this category problems are ranged from those which are highly specific to those which are so general as to underlie the solution of many other problems.

It is, for example, a neat problem to ask the relation between oxygen intake and carbon dioxide output under a given set of activity conditions in man. It is not a neat problem to track down the factors that determine how birds or fish are guided in their migrations. It is a specific problem to determine what colors in a given list a person with a certain visual defect can distinguish; but it is a broad problem to study some mechanisms underlying visual behavior.

We have said that problems of the second class are those which can be answered by experimentation and the rules of logic. Problems of this class pertain to concepts and events that can be defined by a *set of operations*. These operations, rather than some abstract concept which has arisen in the experimenter's mind as something that is ideal or "ought to be," are taken as a starting point.[1] The operations may be perfectly arbitrary, but nevertheless they define just what is meant. Units and other items in physics are defined in terms of an implied set of operations. A foot thus becomes a distance or length coinciding with an arbitrarily chosen unit that is applied. The operation that is implied is the application of the arbitrary unit. Electrical resistance, electromotive force, electric current, etc., are concepts that are derived from the use of a set of operations and the results accruing from them.

[1] Thinking and the communication of thought require the use of words. Words imply definitions. Often words imply the existence of fixed entities. But words may apply to concepts that do not pertain to fixed entities or things. They apply simply to kinds of events. When these events are sharply defined and are repeatable under control, we often speak of them as operations. Putting the matter the other way around, operations can be labeled or named, and so can the results of operations. Nouns, then, are the names not only of things, but also of acts and of the results of acts. Even though there is considerable arbitrariness to the products of operations, the idea of thinking and defining in terms of operations has proved indispensable in science as we know it.

The same principle, that of *operationism,* holds for behavioral events in psychology. One can define any trait, or aspect of behavior, in terms of what is involved in identifying or measuring it. The limitation is that the definitions so derived should bear a systematic relation to each other. This requirement makes it apparent that operationism, if it is to mean much in science, is not a blanket license to universal arbitrariness and chaos. It is meant to be a means toward system building rather than the reverse. All too often this requirement is not obeyed, and the experimenter's concept of the behavior of the human individual is little, if any, better than a hodgepodge.

QUESTIONS

1. Name three or four operations which Bentley listed as being sometimes called experimentation.
2. What is the objective of experimentation?
3. What are the two kinds of assumptions involved in every experiment?
4. What are suppressed premises?
5. What are axioms?
6. Which of the two kinds of assumptions is more likely not to be recognized by the would-be experimenter?
7. Should the experimenter set out to find confirmation of his hypotheses?
8. Give an example of a common or well-known type of experimentation and state some of the major presuppositions.
9. Cite an example of experimentation wherein you disagree with the experimenter upon the validity of the presuppositions involved.
10. Do you agree that the two kinds of assumptions are the basic determiners of the experimental results obtained in any investigation? Justify your position.

CHAPTER 4

RELATING AND MEASURING

Relating and measuring are essential to science. Accordingly they are prominent in experimentation.

In certain areas of science, it is taken for granted that events or phenomena are measurable. In such cases, questions center on the selection of a detector for use in the measurement process, the sensitivity of measurement, and the control of the conditions which produce the effects to be measured. But in biological science, particularly in psychology, there has been, up to now, some doubt as to the measurability of many phenomena. This doubt has brought into play the question of the nature of measurement itself. Hence definition of measurement is a subject of primary concern. Numerous definitions of measurement, from one-word synonyms to formal and extended statements, are possible. Each dictionary definition implies something a little different from the other, and each omits something that probably ought to be included. Among the dictionary terms are the following: to allot; to ascertain the extent, size, or volume of; to mark out; to estimate; to determine by rule or standard.

A definition which will serve as a basis for consideration is the following: *Measurement is the assignment of numbers to events, objects, or properties according to certain rules.* The rules determine how these numbers relate to each other and thus form a scheme or scale. There are four essential kinds of scales, and these will be discussed one by one. Each of these types of scales constitutes a unitary framework. This framework is an understood and predetermined set of relationships between the items in it. It is by virtue of this set of relations that each item possesses a meaning. The fixity of the rules provides the fixity of relationships, and the concept of measurement derives from the

ability to determine given relations of the specific to the general in any case at any time. A unit, such as a foot or yard, does not derive its full meaning by virtue of its own existence but by virtue of a number system with already-known properties to which it belongs.

1. *The nominal scale.* In nominal scales, numerals are assigned only as labels. Their function is such that letters, words, or any other class of symbols would serve just as well. The assigned nominal numbers bear no necessary logical relation to each other or to the total number of assignments constituting the "scale." The numbers serve only the function of *denoting*.

For example, the numbers of two employees known as 116 and 58 respectively do not necessarily bear any logical relationships to each other. They do not represent literal ages of the employees or their months of service. One employee is not twice as old as the other or twice as long in the company's service, nor has he twice as much education as the other, even though the numbers bear a relation of 2 to 1.

Assigned nominal numbers may be used as a *code*, in which case they serve as messages. The number 63 placed on an examination paper by the instructor may mean that the student must be much neater the next time, else his paper will be thrown in the wastebasket.

The numbering in a nominal scale may apply to individuals or it may apply to groups or classes. If the former, we shall call it type A, and if the latter, type B. In type B, it is obvious that several items or individuals would be assigned the same numeral. Actually, the giving of different numerals to different individuals is only a special case of numbering classes in which each class contains but one individual.

The only thing that can be done statistically with a scale of type A is to sum the number of individuals in the scale. In type B, the most numerous class, called the *mode*, may be stated, since the classes may vary in the number of individuals they contain.

2. *The ordinal scale.* The operation of rank-ordering gives rise to the ordinal scale for which ordinal numbers first, second, third, etc., are used. Such scales can be treated mathematically

in any way that preserves the original order and does not imply characteristics not inherent in it. In the strictest use of the ordinal scale, *means* ("averages") and *standard deviations* should not be used. These imply a knowledge of properties of the scale beyond the rank-order of the data. Since the interval represented between ranks is not necessarily the same in all cases within the scale, the *mean* cannot be legitimately obtained. The *median* rank can be computed, however, since it is merely the middle rank without regard to rank intervals. In Table 3 it is rank 9.

To illustrate an ordinal scale, the following example is given. In a class in beginning psychology, 27 students scored points as follows, on an examination. From the raw scores, ranks were assigned, as indicated in column 3 of Table 3.

In the example, two characteristics of ordinal scales are immediately observable. First, it is apparent that ranking is not synonymous with fixing equal-sized intervals between classes. For example, the actual distance between classes (ranks) 1 and 2 is four points, whereas between 3 and 4 it is nine points. Although we know this from inspection of the raw-grade points in column 1, there is nothing inherent in the individual items in column 2 (number of students) to indicate this. All we know is that since rank 1 is made to represent the greatest number of points on the examination, rank 2 represents a just lesser score, and rank 3 a score just less than rank 2. Actually, as far as the properties of the scale are concerned, the situation could be reversed, so that rank 1 would represent the lowest number of points rather than the highest.

Second, it is apparent that each class (rank) is not equally populated. There is nothing in the ordinal scale to affirm or deny the inequality of the populations in the ranks or to tell anything about the number of cases in them. In some instances in the example more than one person obtained the same number of points, hence they are of the same rank.

Events or properties which can be adequately denoted by nominal or ordinal numerals are nonadditive. They cannot be counted. If our definition of measurement is to require *counting* as an operation, nominal and ordinal scales are not to be con-

sidered scales of measurement. It is certainly quite likely that most individuals would not consider the nominal scale a form of measurement, whereas a number of persons would be willing to

TABLE 3

Number of points	Number of students	Ordinal scale	Size of interval in points
120	1	1	4
116	1	2	7
109	1	3	9
100	2	4	1
99	1	5	1
98	2	6	6
92	5	7	2
90	3	8	6
84	2	9	1
83	1	10	3
80	2	11	3
77	1	12	5
72	1	13	2
70	1	14	5
65	1	15	1
64	1	16	4
60	1	17	

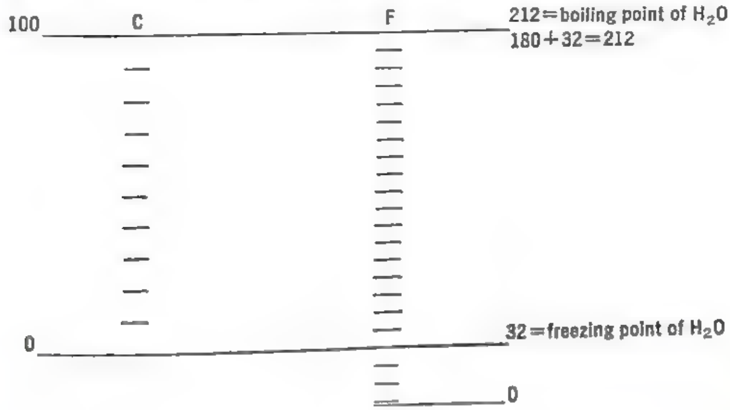
include ordinal scales. Accordingly, the original definition which stated that measurement is the assignment of numerals according to rule is too inclusive.

Supposedly it is not the use of number in and of itself (as in a nominal scale) that would seem to provide the experimenter with the real tools of measurement, but rather the existence of a logical system in which certain defined operations can be performed with these numbers. The ordinal scale is more adequate in this respect than the nominal. We come next to a scale which will provide for a still greater number and variety of operations.

3. *The interval scale.* This scale implies rank-order, and in addition implies equality of intervals between the classes (ranks). By use of this scale, it is possible to begin to "quantify." Cardinal numbers are used in their proper way, and the scale possesses almost every property permitting the employment of common statistical operations. The one property of the scale that is notably absent is the *true zero point.* The zero point used in the scale is one simply of convenience or convention. The scale form is not distorted by the addition of a constant. In fact, the use of a constant is part of the procedure of shifting from one interval scale to another, a way of converting values on one scale to another. Two of the best examples of interval scales are the Fahrenheit and centigrade temperature systems. In both cases, equal intervals of temperature are constructed from equal volume changes in the expansion of some substance such as mercury. The zero points of the two scales are arbitrary and different. To transform a numerical value from one scale to the other, an equation of the form $x' = ax + b$ is used. To use the equation in this case, let x' be the centigrade temperature, and x the Fahrenheit; a, the conversion factor (the relationship of the size of the units in the two scales); and b, the difference in the position of the two zero points. Thus a temperature of 30°C. is equivalent to a temperature of 86°F. ($30 = 9/5\,x + 32$ or 86). Within each scale, the size of the intervals is constant. A degree at the higher levels of temperature is the same as at the lower levels. To this extent the units are additive.

4. *The ratio scale.* As was pointed out, the scales already described are not usable for *counting.* The numbers in the nominal scale label items simply in the way in which letters and other symbols do. The numbers in the ordinal scale signify more-than and less-than relationships. While not providing

for any kind of counting, the scale constitutes a kind of relating process. The numbers in the interval scale are used not only for ranking but also for spacing the relationships denoted.



Scales may go a step farther and provide for a true zero and for dealing with opposites. Scales for doing this are called ratio scales. Ratio scales are used to count events, things, and distributable properties. Here again cardinal numerals are used. In this scale all the properties of the cardinal number system are usable. The ratio scale simply *is* the cardinal number system. In ratio scales, operations for determining equality, rank-order, equality of interval, and ratios are possible. With the ratio scale, all the operations possible with the preceding scales plus those unique to it are possible.

If the classes or items in a scale are actually additive, they will satisfy certain operational criteria. Some of these are so obvious that mention of them might seem unnecessary. But actually it is well to call attention to them.

1. The first of these criteria is that of *asymmetry*. A_1 must be greater than class A_2, and A_2 must be greater than A_3. It might seem better to label the classes in the opposite order and say that A_1 must be smaller than A_2 and A_2 must be smaller than A_3.

2. This asymmetrical relation between classes in the scale must be *transitive*. If A_1 is less than A_2, and A_2 is less than A_3, then A_1 is less than A_3. Although this seems obvious, it need not be. For example, object A_1 when lifted provides for a cer-

tain impression of heaviness. A_2 seems heavier than A_1, but A_3 might not seem any heavier than A_1 than A_2 has seemed, although when A_2 and A_3 are compared, A_3 seems the heavier.

3. The effect of adding two classes together must be the same as the effect of a third class which is equal to their sum. This rule is not always obeyed in psychological experiments.

If an object is placed in a subject's hand, and then a second equally heavy object is added to this, the effect will not be the same as if an object as heavy as the sum of the two were placed in the subject's hand at once. This is but one way of saying that the properties of two objects do not add under all conditions. In the first place, a weight twice as heavy does not feel twice as heavy; and in the second place, it does not feel the same when applied in two installments.

4. The order of adding classes must be immaterial. That is, if A_1 can be added to A_2 to equal A_3, then one should be able to add A_2 to A_1 to equal A_3. This also is violated in many psychological situations.

5. Another possibility that must be satisfied if classes are truly additive is that if A_1 and A_2 taken together and added to A_3 amount to A_4, then A_1 added to the combination of A_2 and A_3 must equal A_4.

There are other criteria to be satisfied, but for our purposes they may be omitted. These five plus two rules regarding opposites will suffice.

If the property in question (x) has an opposite ($\bar{x}$) which can be called the negative of x, the following must be true. The group of objects must contain a hypothetical class, a "neutral" member of A_0 which does not excite the detector of the property. Hence, if A_0 is included with A_1, the result will be the same as if only A_1 were present. If the group contains an object activating the detector in one manner, it must contain A_{-1} affecting the detector in the opposite manner. If these opposite effects are equal, then A_1 plus A_{-1} will not excite the detector at all.

In addition to the classes in a ratio scale being additive as just outlined, the scale has a true zero. All kinds of statistical operations are applicable to such a scale. To convert a value on one ratio scale to a value in another scale, one need only to

multiply the value by a constant. Logarithmic transformations are also permissible in ratio scales.

The distinctions between the ratio scale and all other scales should be kept in mind. The ratio scale, as was already stated, provides for *counting*. It is true that certain definitions of measuring require the operation of counting. Psychology must, therefore, take stock to see whether there are phenomena within its field which can be subjected to the process of true counting. If so, scales should be constructed to demonstrate this. Two of these scales that have already been constructed will be described in Chaps. 8 and 9. A ratio scale has also been constructed for lifted weights. Perhaps other ratio scales will be constructed, and when this is done, psychology will come much nearer to demonstrating to physicists and others that there are psychic phenomena that are measured.

QUESTIONS

1. Distinguish between relating and measuring.
2. Do you believe that all four scales mentioned in this chapter should be thought of as scales of measurement?
3. What is a unit?
4. Do all scientific experiments necessarily involve measurement?
5. Do all quantitative experiments involve counting?
6. What are the characteristics of the nominal scale?
7. What are the main characteristics of the ordinal scale?
8. What are the main characteristics of the interval scale?
9. Briefly, what is meant by additivity?
10. Why cannot averages and standard deviations be used in dealing with the ordinal scale?

PART III

PSYCHOPHYSICAL METHODS

CHAPTER 5

THE METHODS OF AVERAGE ERROR AND LIMITS

In answering questions experimentally, a number of essential procedures are possible. When some of these procedures were reduced to standardized forms, what we know as *psychophysical methods* came into being. They were so named because some of the developers of the methods believed they were using them to discover the relations between mind and body.

The object of this chapter is to describe two of the several psychophysical methods. But before doing so, it is appropriate to point out the more general operations that are possible for subjects or observers to perform.

First is that of *naming*. The items named may be things, qualities, traits, etc. Naming is an essential operation, for to name is to identify and to imply distinctions between things or properties. This procedure is one's first task in any field that has either not been worked over or has been improperly dealt with. In classifying, experimenters perform the operation of naming. In experiments, subjects may be called upon to name.

A second operation, that of *defining*, is an extension of the process of naming. While naming may or may not imply rigid or formal distinctions, defining certainly does. Naming and defining need not, in themselves, be formal experimental procedures such as are involved in collecting data, although they sometimes are just that.

The next operation to be pointed out is that of *rating* or *ordering*. Objects or properties are placed in a scale of *more or less*, without regard to actual amounts. This is a common performance required of a subject.

Quantitatively the most precise operation, however, is that of comparing. Comparisons may involve either the successive

or the simultaneous presentation of the items to be judged. Accordingly, there are simultaneous and successive comparisons. The former are much more precise.

Not all comparisons are of the same sort. The most exact are those of *equating* or *matching* two items. One item is adjusted to appear *equal* to the other, or one is used as a standard and another item that will *match* it is chosen from a group. A second form of comparing is that of *fractionating.* This may be done by having the subject choose a value one-half (or some other fraction) as great as another item used as a standard. This procedure applies equally well to adjusting a variable or to choosing an actual item of fixed value from a group supplied for that purpose.

A third form of comparing is that of *multiplying.* A value is required here that is some multiple of the value of a given standard. Here again a variable having this possibility may be used, or an item may be chosen out of a group of given items.

A fourth form of comparing is that of *sectioning or bisecting.* In such cases two standards are supplied—two terminal values; one a low, the other a high. The chosen value or item in bisecting is perceived to lie midway between the two standards.

Still another form of so-called comparison is that of deciding which of two items is preferable in some respect. This is done when a group of items is involved. For example, a number of colors may be studied from the standpoint of human preference. Two items at a time are thus compared until every item is paired off with every other one once. It would be better to call this operation that of *preferring,* although comparison is supposedly involved in the act. We shall come upon this procedure again in Chap. 7 under the section on paired comparisons.

Still another operation is that of evaluating single items. It is called the operation of *absolute judgment.*

A tabulation of the operations just discussed is shown in Table 4. The items in the second column, however, are by no means mutually exclusive.

As was said, the formal procedures developed and used at the time psychology became definitely experimental (see Chap. 1) were called psychophysical methods. The study and general

employment of these methods is now known as psychophysics. We shall deal with the major psychophysical methods one at a time.

TABLE 4

Procedure	*Function*
A. Naming	Denoting and identifying
B. Defining (relating) verbally	Identifying and classifying
C. Rating and ordering (putting in sequence)	Ordering or arranging
D. Comparing	Quantifying or measuring
1. Equating and matching	
2. Fractionating	
3. Multiplying	
4. Bisecting—trisecting (sectioning)	
5. Preferring	
E. Absolute judgment	Evaluating single stimuli

The Method of Average Error. The method of average error, also called the method of *reproduction*, consists in making a large number of readings wherein the observer himself manipulates the variable. For example, an observer may be presented with two lighted rectangles, in a darkroom, that he is called upon to match in brightness. The brightness of one is fixed and that of the other is variable and is adjusted by turning a knob. The observer begins by having the variable rectangle too bright or too dark. He adjusts its brightness level up and/or down until he reaches a satisfactory match between the two rectangles. This constitutes one trial, and, as was said in the beginning, a great number of trials are made. This is so that the results may be amenable to statistical treatment. The arithmetic mean of all the readings will differ at least slightly from the brightness value of the standard rectangle. Likewise, the readings tend to differ from each other. When many of them are made, they will tend to cluster rather symmetrically around the mean. This is indicated in Fig. 1.

In this figure, *S* and *V* represent two bright rectangles in a darkroom. *S* is of fixed brightness, let us say, 100 candles per square foot (c./ft.^{2}). *V* is variable, and the aim is to match *S* in brightness. The average value of many trials is, let us say, 103 c./ft.^{2} All readings, however, are not 103, but scatter from a value somewhat below 100 (in this case, 96) to a value above 103 (in this case, 112) c./ft.^{2} Although several readings may turn

out to be of the same value, no two values need represent the same number of readings. The height of curve D, at any point, indicates the numbers of readings at given value along the brightness axis B.

The difference between the fixed value of rectangle S and the mean value of readings for V is called the "error" of overestimation or of underestimation, as the case may be. In either case, it is the *constant error.* This tendency to vary from an exact

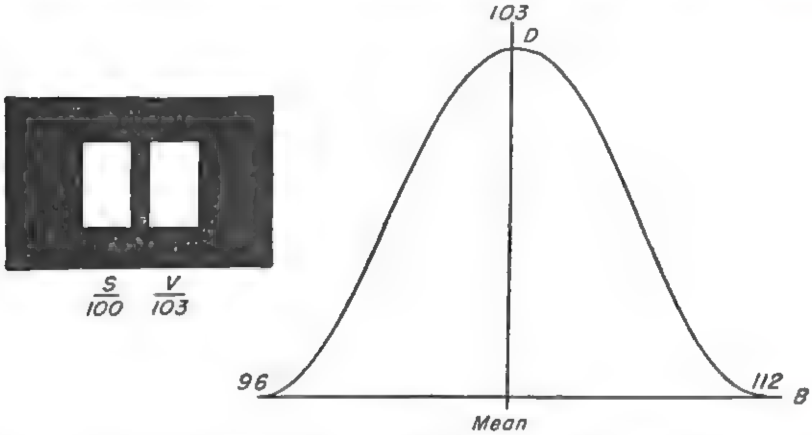


FIG. 1. At left is shown a test-field in which two luminous rectangular areas are to be compared in brightness by adjusting the intensity of area V. At right is a curve showing a hypothetical distribution of the values of a series of readings when areas V and S are made to match subjectively. Values are in candles per square foot (c./ft.^{2}).

match with the brightness of S is one type of tendency, whereas the tendency to vary from trial to trial in the values of the readings is a very different type of tendency. The latter "error" is called the *variable error.*

The discrepancies in measurement occur from the fact that subjective behavior (perception) is such that two physical qualities may differ by a finite amount before a difference is observable. In other words, the difference must reach a *threshold* amount before being detectable. The measurement of thresholds is one of the major objectives in psychophysical procedures. The value of the second stimulus as it differs from some standard is the *threshold value.* If the standard stimulus has a finite value (above zero), then the stimulus that reaches a threshold value is

spoken of as the *differential threshold*. The difference in value between the two stimuli is spoken of as the *difference limen*.

If the standard is zero, then the smallest stimulus that will evoke a response is called the *absolute threshold*. The value also marks off an interval between it and zero, which is called the *absolute limen*. Figure 2 will clarify these points.

The error in any particular reading then consists of two components, the constant and the variable. The variability is also

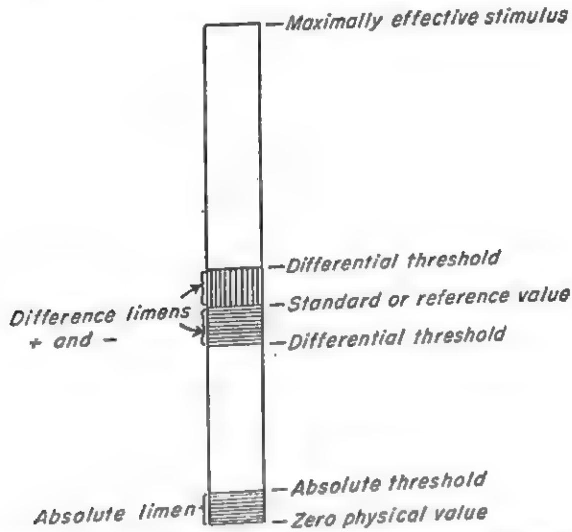


FIG. 2. Schema to illustrate the relationships between the various concepts involved in obtaining thresholds by the method of average error.

spoken of as *sensitivity*, since it represents the stability of judgment or the ability to repeat a given value on successive readings. The other error (constant) represents a very different factor. At times this may become great, and when so, the situation used to be looked upon as illusory.

The Method of Limits. This has been sometimes called the method of *minimal change*. In this procedure, the experimenter rather than the subject makes the adjustments of the variable. Let us say we use the very same setup as for illustrating the previous psychophysical method. The experimenter begins in one trial definitely below the brightness of the standard rectangle and slowly increases the intensity of the variable rectangle. When the two rectangles appear to be equally bright, the observer

out to be of the same value, no two values need represent the same number of readings. The height of curve *D*, at any point, indicates the numbers of readings at given value along the brightness axis *B*.

The difference between the fixed value of rectangle *S* and the mean value of readings for *V* is called the "error" of overestimation or of underestimation, as the case may be. In either case, it is the *constant error*. This tendency to vary from an exact

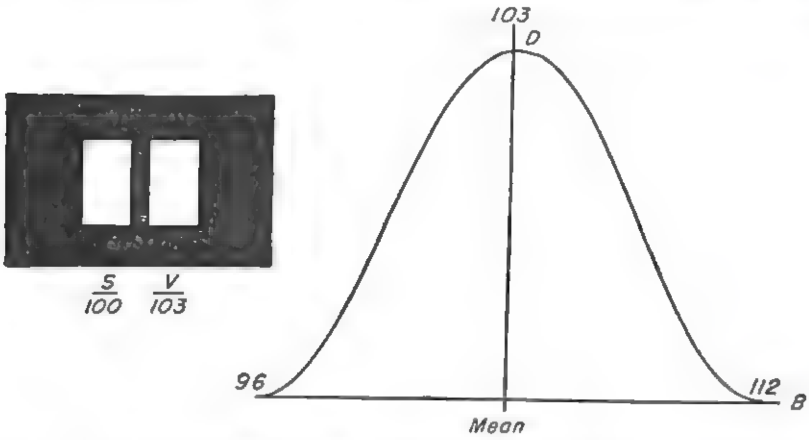


FIG. 1. At left is shown a test-field in which two luminous rectangular areas are to be compared in brightness by adjusting the intensity of area *V*. At right is a curve showing a hypothetical distribution of the values of a series of readings when areas *V* and *S* are made to match subjectively. Values are in candles per square foot (c./ft.^{2}).

match with the brightness of *S* is one type of tendency, whereas the tendency to vary from trial to trial in the values of the readings is a very different type of tendency. The latter "error" is called the *variable error*.

The discrepancies in measurement occur from the fact that subjective behavior (perception) is such that two physical qualities may differ by a finite amount before a difference is observable. In other words, the difference must reach a *threshold* amount before being detectable. The measurement of thresholds is one of the major objectives in psychophysical procedures. The value of the second stimulus as it differs from some standard is the *threshold value*. If the standard stimulus has a finite value (above zero), then the stimulus that reaches a threshold value is

spoken of as the *differential threshold.* The difference in value between the two stimuli is spoken of as the *difference limen.*

If the standard is zero, then the smallest stimulus that will evoke a response is called the *absolute threshold.* The value also marks off an interval between it and zero, which is called the *absolute limen.* Figure 2 will clarify these points.

The error in any particular reading then consists of two components, the constant and the variable. The variability is also

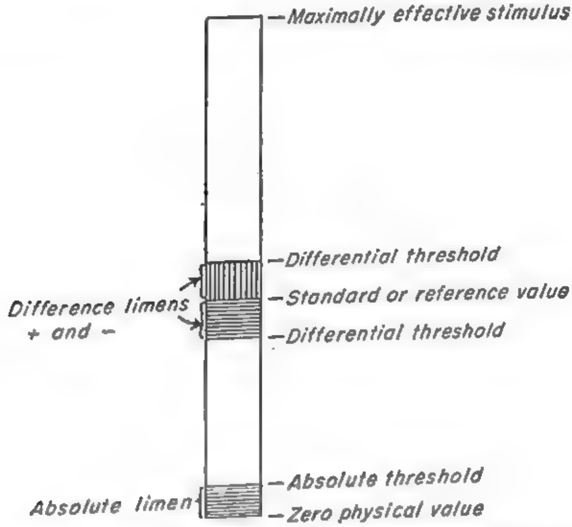


FIG. 2. Schema to illustrate the relationships between the various concepts involved in obtaining thresholds by the method of average error.

spoken of as *sensitivity*, since it represents the stability of judgment or the ability to repeat a given value on successive readings. The other error (constant) represents a very different factor. At times this may become great, and when so, the situation used to be looked upon as illusory.

The Method of Limits. This has been sometimes called the method of *minimal change.* In this procedure, the experimenter rather than the subject makes the adjustments of the variable. Let us say we use the very same setup as for illustrating the previous psychophysical method. The experimenter begins in one trial definitely below the brightness of the standard rectangle and slowly increases the intensity of the variable rectangle. When the two rectangles appear to be equally bright, the observer

signals the experimenter. In the trial to follow, the experimenter begins at a brightness level definitely above that of the standard and lowers the intensity until the observer signals him he is satisfied. If the readings obtained in the ascending trials are averaged separately from those of the descending trials, two "errors" are designated, the ascending threshold and the descending threshold. The differences between the standard and the ascending and descending thresholds are, of course, called the ascending and descending limens, respectively.

Instead of a gradual increase or decrease in the brightness of the variable, a stepwise shift in each direction may be made.

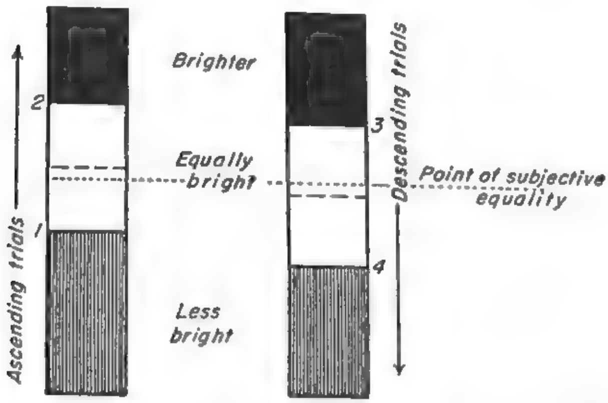


Fig. 3. Schema to indicate the relationship between the various concepts dealt with in obtaining thresholds, etc., by the method of limits.

In such cases, a response of the observer will be made to each step up or down in brightness. While the observer still sees the variable as less bright than the standard, he says so. When he sees it brighter, he says so. Thus as the brightness of the variable is being increased from a level definitely below the brightness of the standard, the observer will be indicating it to be less bright. Finally, as the stepwise adjustment upward continues, he will see the variable rectangle as brighter. The experimenter uses the mean between this value and the preceding one as the value to note down. The same principle holds true as brightness is being reduced.

While, in what has just been said, two categories of response have been allowed, it is possible to give instructions to the

observer for three. The observer may be instructed to respond by saying either "less bright," "equally bright," or "brighter." In this case, as the two rectangles *approach* equality, the response will be that of "equally bright," and will continue to be until the two rectangles pass beyond being equally bright. Then the response will begin to be "brighter" or "less bright," as the case may be.

Figure 3 illustrates the method of limits using three categories of judgment. It is apparent that there will be four boundary conditions, two (1 and 2) in the ascending trials, and two (3 and 4) in the descending trials. The point of subjective equality of the variable and standard is taken as the mean of the four readings. The procedure may also employ stepwise presentations, and then the principle previously explained for noting values in such cases is to be followed.

QUESTIONS

1. What is a psychophysical method?
2. Are all psychophysical methods quantitative?
3. What are the various general operations that a subject may be called upon to perform?
4. Why is the operation of "absolute judgment" so called?
5. Is the operation of "preferring" a means of true measurement?
6. What are the essential similarities of the methods of limits and of average error?
7. What are the essential dissimilarities of the methods of limits and of average error?
8. Can you suggest experiments in which one method would be more suitable than the other?
9. What is the difference between a threshold and a limen as defined in this chapter?
10. What is the difference between a constant error and a differential limen?

CHAPTER 6

THE METHOD OF CONSTANT STIMULI

In the previous chapter two of the standard psychophysical methods were described. In this chapter another method is to be presented. It is the *method of constant stimuli*, also called the constant method.

In planning to use this method, the experimenter chooses a number of fixed values ("constant stimuli") of the variable, one-half of which lie above and one-half below the standard. In so doing, he must have sufficient preliminary knowledge of the character of the observer's response to know how great a range to cover. This knowledge he often obtains by a few preliminary trials before starting the formal part of the experiment.

The method of constant stimuli satisfies two preferred conditions in performing the operation of comparison. The first of these is the prevention of the subject's knowing which is standard and which is the variable stimulus. All the subject is asked to do is to judge which of the two stimulus objects possesses more of the property in question. This property, which we may designate as "*x*-ness," may be that of length, width, intensity, loudness, brightness, redness, smoothness, or any other one of a very indefinite number of possible qualities. The other condition which the method of constant stimuli satisfies is that of not making known which of the two stimuli is being increased or decreased.

The method of constant stimuli is based upon certain statistical assumptions, the first being that the greater the actual physical difference between the stimuli used, the greater will be the number of trials out of 100 in which the difference will be perceptible. The converse, of course, is also true. Naturally the theoretical range between maximum and minimum percent-

ages in correct judgment is from 0 to 100. In actual practice it would not be expected that the observer would be incorrect in his judgments all the time, if the range of values chosen by the experimenter is appropriate. Even by merely guessing, the subject would be correct part of the time. If there are only two possible choices of response, he would run the chance of being correct in 50 per cent of the trials. The range of fixed values of the variable must be small enough for the resulting difference between the extremes and the standard to preclude the observer being correct in every trial. If the differences were large enough for the observer's judgments to be correct in every trial, the information obtained by the use of such stimuli would be statistically valueless. The experimenter would not know how much the differences between the chosen values and the standard would need to be reduced before the percentage would drop just below 100.

The data that are obtained by the method of constant stimuli express the relation between correct judgments and the magnitude of difference in some physical property of the two items presented (standard and variable). Therefore the graph representing the findings is a plot of these two factors. In Fig. 4 is given an illustration of a set of such relationships. The experiment consisted of presenting to the subject two circular areas (disks) of light. The means for doing this was provided by a partitioned lamp house with two opal-glass windows, the disks. Each compartment was lighted separately. The intensity of the lights could be varied separately in a controlled fashion. The subject was asked merely to state which of the two disks appeared brighter. Since the experimenter knew the amount of difference in the intensity of the lights, he obtained data by which the relation between magnitude of intensity difference and correct judgments could be determined. Actually one of the two lights was kept at a constant level and served as the standard. This standard value was shifted from one to the other in chance order. If the experimenter had not cared to prevent the subject's knowing which of the two disks was the standard, he could have set one of the two disks at a fixed intensity value and left it there throughout the experiment.

Possibly in an experiment of this sort it would not make too much difference in the final outcome to conduct the experiment in this way. At least the student should understand that these two alternatives are possible, for in some experiments it might make an essential difference in the subjects' judgments to have the standard always on the left or always on the right.

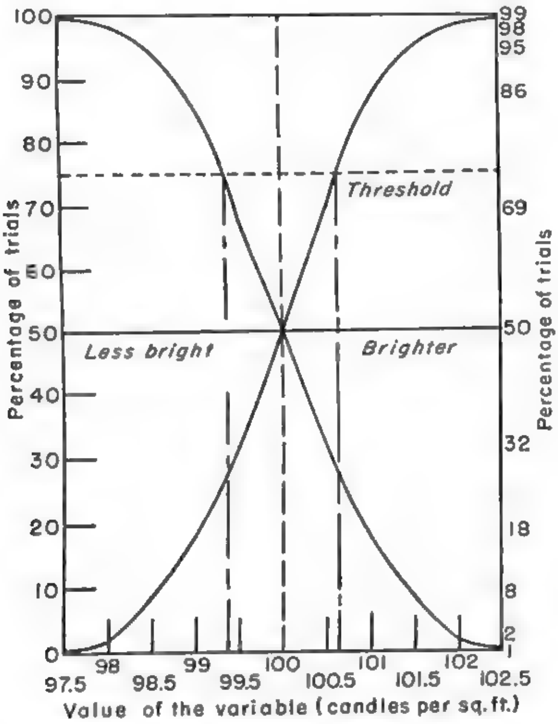


FIG. 4. A graph to show the relation between the value of the stimulus variable (in relation to the standard) and the percentage probability of making a correct judgment when one of two responses is allowed. The response must be either "greater than" or "less than" the standard.

The Graph. In the graph in Fig. 4 the relationship between the intensities of two disks of light and the correctness of judgment is plotted. The actual intensities of the disks vary from 97.5 candles per square foot (c./ft.^{2}) to 102.5 c./ft.^{2}, in steps of 0.5 c./ft.^{2} The standard value was 100 c./ft.^{2} Naturally the disk that happened to be less intense than this was judged brighter part of the time (in part of the trials), but was judged

less bright in most cases. The values and the judgments shown in the graph are also given in Table 5.

TABLE 5

	Judged brighter, per cent of trials	Judged less bright, per cent of trials
97.5	1	99
98.0	2	98
98.5	8	95
99.0	18	86
99.5	32	69
100.0	50	50
100.5	69	32
101.0	86	18
101.5	95	8
102.0	98	2
102.5	99	1

Categories of Judgment. If the subject is allowed to give one of two or three kinds of response, the results will be as follows. For example, the subject may be instructed to respond by saying one is "brighter" than the other, or that they are "equal." Of course, if the experimenter instructs the subject always to use one of the two lights as the reference (standard), then the subject may be permitted to respond by one of three answers, "brighter," "less bright," or "equal." The graph in Fig. 5 indicates the type of outcome ensuing from instructions permitting three categories. There will be three curves in the graph instead of two. The third curve represents the percentage of equality judgments for each of the values of the variable. Naturally the number of equality judgments will be greater, the smaller the actual intensity difference between the two disks. Nevertheless, a few equality judgments are likely to occur even when the differences between the disks are extreme. Since there are three responses possible, the percentage of equality judgments, by pure chance, would be in the neighborhood of 33.3 per cent when both disks are equally intense. In case the three categories of judgment ("equal," "brighter," "less bright") would turn out to give equal percentage at this point, then the three curves

would intersect at 33.3 per cent. As the curves are drawn, however, the equal category represents 55 per cent when both disks are set at 100 c./ft.[2] The "less bright" and "brighter" curves together then account for approximately 45 per cent of the judgments. Each of these two choices, naturally, were made fewer times than if no judgment of equality were allowed.

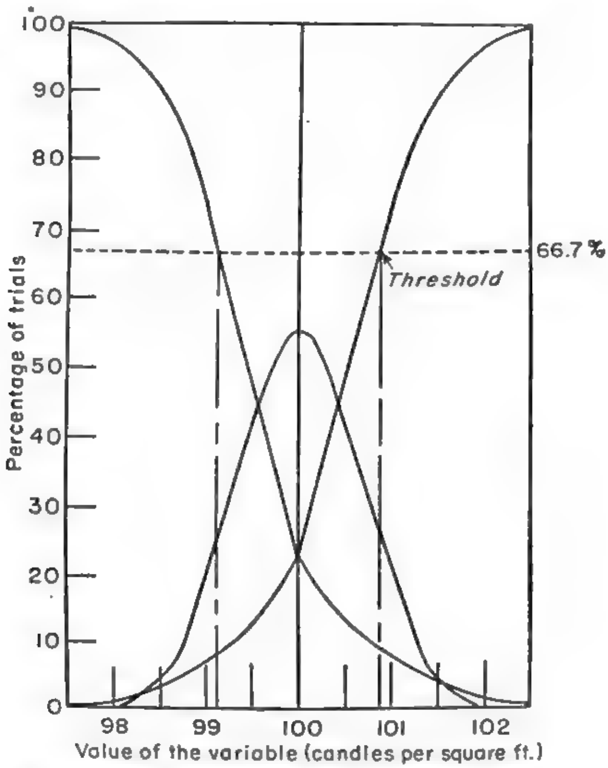


FIG. 5. A graph to show the relation between the value of the stimulus variable (in relation to the standard) and the percentage probability of making a correct judgment when one of three responses is allowed. The responses in this case are "equal to," "greater than," or "less than" the standard.

The question of whether two or whether three categories of judgment are preferable has been debated among authorities. Be that as it may, it is obvious that if equality judgments are allowed, trials will be put into this category when the subject is not quite sure. The two disks will be judged equal when no difference is detected. On the other hand, when only two judg-

ments ("brighter" and "less bright") are allowed, a sharper kind of discrimination is called for. It is therefore not nearly so artificial to confine the judgments to two categories as it may seem. Actually, in all but 9 per cent of the cases there is a finite difference in physical intensity in the stimuli. Consequently, seldom is the subject asked to "perceive" a difference when it does not exist.

Threshold. Our next problem has to do with the value to be chosen as the threshold when experimentation is conducted in the way just described.

The definition that applies in general applies here. It makes the threshold a statistical affair, taking into account the operation of chance under the given conditions. The threshold is the physical value of the stimulus halfway between one representing chance probability and 100 per cent correctness (correctness in every trial). With two alternatives of response where one or the other is always correct, pure guessing would yield a correct response in 50 per cent of the trials. To satisfy the definition, the 75 per cent level on the curves would then represent threshold. This level is shown in Fig. 4 by a horizontal broken line extending to the curves, from which perpendiculars are projected to the base line to indicate stimulus values. These values need not be any of those actually used as the "constant stimuli." In fact, the threshold values generally are not any one of those actually used.

If three responses instead of two are provided for, chance correctness would be 33.3 per cent. In this case threshold would be represented by the stimulus value corresponding to 66.7 per cent. To show what this value is, broken lines are located in Fig. 5.

A comparison of the two figures will show that the thresholds obtained with three categories of response are not the same as those obtained with two categories of response. The latter appear to be smaller, and this difference is in keeping with the earlier suggestion that the two categories make for greater precision of judgment.

In both the case of the two categories of judgment and of the three, the point at which the "brighter" and "less bright"

curves intersect is called the point of *subjective equality*. The difference between this value and the actual value of the standard stimulus is called the *constant error*. In the curves shown in Figs. 4 and 5, the constant error is zero. Generally the constant error has some finite value. The difference in value between the point of subjective equality and the lower threshold value is called the *lower limen*. The difference between the point of subjective equality and the upper threshold is called the *upper limen*. The two do not have to be equal in actual practice.

When only two categories of judgment are allowed, the curves intersect at the 50 per cent level. When three categories are allowed, the curves intersect below the 50 per cent level, actually close to the 33.3 per cent level. If perpendiculars are erected to touch the curves at the 50 per cent level, an interval is thus marked off on the stimulus axis (abscissa). This is called the *interval of uncertainty*. Some workers arbitrarily call the midpoint in the interval of uncertainty the point of subjective equality and define the thresholds as the values when the perpendiculars strike the stimulus axis. When this rule is followed, the limens are the values represented by one-half the interval of uncertainty.

QUESTIONS

1. What is one of the principal advantages of the method of constant stimuli over the method of limits and the method of average error?

2. What difference in results may be expected when three categories of response instead of two categories are used?

3. Does the subject or observer need to know which of the two presentations is the standard when the method of constant stimuli is used?

4. Why is one of the fixed stimulus values not so chosen as to provide for correct response in 100 per cent of cases?

5. Why is the threshold, according to the method of constant stimuli, said to be a statistical affair?

6. When using the method of constant stimuli, might other values aside from the threshold be of interest? Give an example, if you can.

7. When using three categories of response, about what percentage of correct responses might be expected when the standard and variable are of the same value?

8. The combined percentage of responses for the three categories of response should total what?

9. In general, what is the most usual shape of the curve describing the distribution of response of "greater than or less than" obtained by the method of constant stimuli?

10. What is the shape of the curve representing the "equal" responses?

CHAPTER 7

THE METHODS OF SINGLE STIMULI AND PAIRED COMPARISONS

The two chief psychophysical methods which remain to be described are the *method of single stimuli* and the *method of paired comparisons*. The *method of single stimuli* is also called the *method of absolute judgment*.

The Method of Single Stimuli. The essential procedure of most psychophysical methods consists in making direct comparisons between two things at a time. These generally are selected attributes or properties of objects. The full procedure consists in comparing one object called the standard and a number of objects, one at a time. According to these methods, the comparison items are either presented *simultaneously* or in quick succession.

According to the method of single stimuli, no such comparison is possible. No formal standard is at any time presented to the observer. The procedure consists simply in the presentation of single items to which the observer is called upon to make some sort of a descriptive response. If it is weight that is being dealt with, he says it is "heavy" or "light." If it is some other attribute, some descriptive term for that attribute constitutes the response. He simply says it is "very black," "slightly black"; "big," "medium-sized," "little"; "warm," "cold," "hot"; "near," "far"; "wide," "narrow"; or whatever he senses it to be in the category with which he is dealing. Obviously the stimulus series consists of a number of items all differing in the amount of the single property in question. This series is given in some kind of a chance order, depending upon what the problem happens to be. The first item presented may be a chance selection from the series. If a number of runs through the series are

made, the initial item may differ for each run. If the procedure itself is under investigation, however, different groups of subjects might be tested in different ways: one group in which the initial items bear a given position in the series, and other groups in which the initial items hold other positions. Another factor in testing the use of this method would be to vary the length of the series used to determine its effect upon the value assigned to given items. The series may contain items uniformly distributed in value, or the population may be quite irrregular in this respect.

If, for example, in size discrimination, the first item presented happens to be judged as small, the remaining items may be so chosen as actually to be larger than the first item presented. Again, the series may lie on either side of the literal value of the first presented item, or the series may consist of members all smaller than the initial item. By some such device as the selection of a series so as to bear some intended relation to the first item, or the first few items, the way in which *value judgments* are developed may be investigated. It is also possible to shift the series of items one or more times as the experimentation progresses, so as to determine the effect of the series as a whole upon the various single items in it.

The data obtained by the method of single stimuli can be handled by the same procedures as used in the method of constant stimuli. When there are two categories in the scale, the means and standard deviation of the transition zone between them can be found.

The method of single stimuli implies a comparison since the responses are given in relative terms, such as "big," "little," and "medium." And we may ask what the standard of comparison is. It would appear that it is the scale formed by the items in the particular series employed. Let any change occur in the series which is presented and it is likely to be reflected in the response to the given item. The method has been called the method of absolute judgment only on the basis of the absence of a concrete, single-comparison standard. Probably, however, judgments here are as nearly absolute in principle as in any other situation.

Let A, B, and C be the items.
$3^2 = 9$ comparisons to be made

TABLE 6

AA	*BA*	*CA*
AB	*BB*	*CB*
AC	*BC*	*CC*

This obviously includes comparing each item with itself. If we eliminate such comparisons, we should have $3(3 - 1)$ or $3 \times 2 = 6$ comparisons.

AB	*BA*	*CA*
AC	*BC*	*CB*

To indicate how this would work with six items, the following is given.

Let A, B, C, D, E, and F be the items.
$6^2 = 36$ comparisons to be made

TABLE 7

$\underline{AA}$	*BA*	*CA*	*DA*	*EA*	*FA*
AB	$\underline{BB}$	*CB*	*DB*	*EB*	*FB*
AC	*BC*	$\underline{CC}$	*DC*	*EC*	*FC*
AD	*BD*	*CD*	$\underline{DD}$	*ED*	*FD*
AE	*BE*	*CE*	*DE*	$\underline{EE}$	*FE*
AF	*BF*	*CF*	*DF*	*EF*	$\underline{FF}$

If we eliminate comparisons of an item with itself, then there should be $6(6 - 1)$ or $6 \times 5 = 30$ comparisons. Excluding the six cases underscored in Table 7 from the 36 comparisons, the remainder is 30.

There is a further requisite involved in the comparison procedure. There must be a chance order so that all of the comparisons involving any given item are not made in succession but are randomly distributed throughout the entire group of trials. Whereas each of the necessary comparisons can be noted on separate small cubes, or small slips of paper, and shaken in a vessel and withdrawn one at a time to indicate a random order, other devices may be employed. One of these is designated in Table 8.

TABLE 8

	A	*B*	*C*	*D*	*E*	*F*
A	(1)	19	7	25	13	31
B	20	(2)	26	8	32	14
C	3	21	(9)	27	15	33
D	22	4	28	(10)	34	16
E	5	23	11	29	(17)	35
F	24	6	30	12	36	(18)

In Table 8 the items *A*, *B*, *C*, etc., are used to head respective columns and to label the rows. To fill in the order of comparisons, one simply writes numbers from 1 to 36, beginning in the upper left corner and zigzagging systematically until all squares are filled in. In Table 9 the actual comparisons arrived at by construction of Table 8 are listed in numerically consecutive order. The result provides one out of many reasonable distributions of trials.

TABLE 9

1. *AA*	7. *AC*	13. *AE*	19. *AB*	25. *AD*	31. *AF*
2. *BB*	8. *BD*	14. *BF*	20. *BA*	26. *BC*	32. *BE*
3. *CA*	9. *CC*	15. *CE*	21. *CB*	27. *CD*	33. *CF*
4. *DB*	10. *DD*	16. *DF*	22. *DA*	28. *DC*	34. *DE*
5. *EA*	11. *EC*	17. *EE*	23. *EB*	29. *ED*	35. *EF*
6. *FB*	12. *FD*	18. *FE*	24. *FA*	30. *FC*	36. *FF*

The comparisons indicated in the foregoing tables include not only those in which, for example, *A* is compared with *B*, but also those in which *B* is compared with *A*. This means that if two square samples of color were being compared, one being red, the other green, in the one trial, red would be presented on the left and green on the right. In the second trial involving these two colors, red would be presented on the right and green on the left.

RED GREEN GREEN RED

When both types of comparison were included, the total number was either n^2 or $n(n - 1)$, depending upon whether the items were compared with themselves or not. If we omit one of the two arrangements, the comparisons are reduced to one-half the original number. This means that there would be only 15 trials. The data obtained from making a series of pair-by-pair comparisons are used for a count of the number of times each of the items was preferred out of the total number of comparisons.

When this information is obtained, the matter of preference can be stated in several ways. The first of these is by use of the *preference score*. This is simply the number of times any given item was given preference. The next is the *relative preference score*. This is obtained by dividing the preference score by the number of items compared, minus one. In terms of formula it would be expressed as:

$$p = P/n - 1$$

where p = relative preference score; n, the number of items, and P the preference score. For example, if there were 15 items and a certain item was preferred six times, its preference score would be 6. Its relative preference score would be $6/(15 - 1)$, or 0.42 +. If the item had been chosen fourteen times, its relative preference score would have been $14/(15 - 1)$ or 1.00. There is also another measure that is often involved in stating results in the method of paired comparisons. It is called the *average preference score*. It is the number of cases minus one, divided by two. For example, for 15 cases the average preference score would be $(15 - 1)/2$ or 7.

As a literal example of the use of the method of paired comparisons, we cite the work of Folgmann on the preference of a group of musicians for the work of the great composers. This list below is given in order of preferences without giving the preference scores for each composer. There were 308 musicians from four of the country's great symphony orchestras who served as subjects (judges), and they were asked to make their preferences, in each case, on the basis of the music and not the composer as a man. The list is as follows:

1. Beethoven
2. Brahms
3. Wagner
4. Mozart
5. Bach
6. Schubert
7. Haydn
8. Debussy
9. Schumann
10. Mendelssohn
11. Tchaikovsky
12. Berlioz
13. Franck
14. Chopin
15. Verdi
16. Stravinsky
17. Grieg
18. MacDowell
19. Herbert

There are additional statistical treatments to which the information from a series of paired comparisons can be subjected. These are too complex for our purposes here. They indicate, however, how reliable the information tends to be.

QUESTIONS

1. Why is the method of single stimuli called the method of absolute judgment?
2. Give an example of a situation in which this method would appropriately be used.
3. What is used as a reference value in the method of single stimuli?
4. Is the reference always the same even during a single series of trials in the method of single stimuli?
5. What factor in the experimental procedure has most to do with the degree of feeling of certainty associated with response?
6. What are value judgments?
7. Are all judgments value judgments in psychophysical experiments? Explain.
8. What psychophysical methods in particular could well be used in social experiments?
9. What operation is performed by the observer in the method of paired comparisons?
10. What scores are used in the method of paired comparisons?

PART IV

THE BROAD FEATURES OF SENSORY EXPERIENCE

CHAPTER 8

THE DEVELOPMENT OF A UNIT OF MEASUREMENT IN HEARING

One of the most significant developments in recent years in the experimental psychology of hearing was the construction of the *sone* scale. This represents, for the first time, true measurement of a psychological magnitude. Some of the early leaders in experimental psychology assumed that the *just noticeable difference* (j.n.d.)[1] was a quantity which had the same value at all levels of stimulation and thus was a true unit. Were this the case, the j.n.d. would have been satisfactory. But the assumption that j.n.d.'s are equivalent in all parts of a stimulus range has always been open to question.

In Chap. 4 on Relating and Measuring, several scales for dealing with experimental data were presented and discussed. One type of scale, the ratio scale, was shown to provide for true measurement. Every operation to which the cardinal number scale is subject is applicable to this scale. It is significant that this type of scale has now been tried in auditory sensation and has been found workable. Before describing the construction of the sone scale, we had better review the various attributes of tones (sounds produced by simple vibration sources).

Tones possess several qualities that can be identified by most

[1] Just noticeable differences are generally taken to be synonymous with difference limens. Preferably speaking, a just noticeable difference is a single nonstatistical event arising as an observation when two stimuli become sufficiently dissimilar. Since the event, though actual, is single (isolated) and not representative, a statistical measure must be substituted. The difference limen is the statistical measure of the just noticeable difference and possesses the predictive value that the j.n.d., strictly speaking, does not possess. In every quarter this distinction has not been made and consistently maintained. In which case the two terms are used interchangeably.

observers. These are the properties that are commonly spoken of in everyday affairs, namely, *loudness*, *pitch*, and *volume*. To these may be added another and less common attribute, *density* (or *brightness*). Furthermore, sounds are experienced as existing for a finite length of time and are heard as existing or "coming from" somewhere. Thus it may be said that they have *duration* and *localization*, or *direction*.

To measure tone is first to select some single one of its attributes and see how it changes as some physical aspect of the sound stimulus is varied. In general, there are two intrinsic features of the physical stimulus that can be varied—its intensity and its vibration frequency. These can be manipulated singly, together, or in opposite directions.

Many investigators in the realm of sound and hearing have chosen the attribute of *loudness* for study. In dealing with this attribute, a scale that would apply to the sound stimulus first had to be developed.

The scale for measuring and describing sound-stimulus intensities is called the *decibel scale*. The unit is the *decibel* and is *defined as ten times the logarithm of the ratio between two energies or powers*. One of these energies is the actual value in question, and the other is a lower limiting value, such as a threshold stimulus. The decibel scale is such that it can be applied to scaling other quantities as well as intensities of sound. When the decibel is used to designate the ratios of two *pressures*, *velocities*, *voltages*, *currents*, etc., the factor is 20 instead of 10. If N represents decibels, then

$$N = 10 \log \frac{E^2}{E_1} = 20 \log \frac{p^2}{p_1}$$

E^2 and p^2 are the magnitudes of the energies and pressures in question, respectively. E_1 and p_1 are the minimal quantities used as references.

The limiting range of common sound sources covers 120 decibels. The intensities of some of the more common sources are shown on the dial in Fig. 6.

Naturally, the decibel scale is not meaningful in terms of one's own experience. Thunder representing 120 decibels (db) is more than twice as loud as conversation, the decibel value of which is

60. Likewise, the din in the average office (40 db) and the sound of a pneumatic drill (80 db) do not seem to have a subjective loudness ratio of 1 to 2. We must remember that the decibel scale does not measure psychological quantities. In fact, although the decibel scale has often been called a *loudness* scale, it is not that but is, rather, a stimulus-intensity scale.

The object of this chapter is to describe how an actual loudness scale was established and to indicate the relation between it and other scales having something to do with the intensity aspect of hearing.

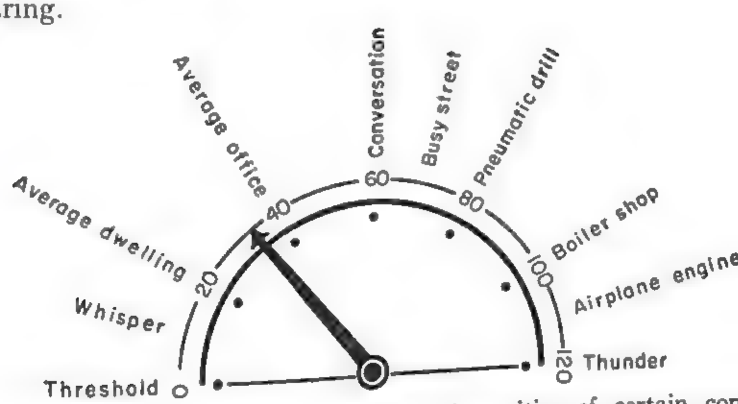


FIG. 6. A diagram showing the sound intensities of certain common sources, according to the decibel scale.

What we want is a scale that is derived by varying the stimulus property which varies one's sensation of loudness. The stimulus property which does this is primarily that of intensity. When stimulus intensity is varied, we want to find those points at which the subject is able to say that one sound is half as loud or is twice as loud as another. If we were to assign the number 10 to the one sound, we should have to assign the number 5 to the sound one-half as loud and 20 to the one twice as loud.

It is obvious that one could not continue to go up or down the scale indefinitely saying that a sound is 1/4, 1/5, 1/6, 1/7, 1/8 . . . $1/n$ as loud as another, or that one is twice, three times, four times, or n times as loud. There is a limit at some point where one ceases to be able to make comparisons. It has been reported that some success has been attained in having observers judge sounds to be 1/10 as loud as a standard.

Owing to the limitations just pointed out, the experimenter shifts the value of the standard up and down the intensity scale and at each point chosen selects another sound which, to a listener, is one-half as loud. These readings can be plotted in relation to the value of the physical intensity of the stimulus measured in decibels.

The two ears are connected neurologically in such a way that a tone presented to one ear sounds half as loud as the same tone presented to the two ears. It is as if some kind of connection existed whereby the effects in the two ears would sum arithmetically. One experimental procedure based upon the fact of summation is to have the observer adjust the intensity of the tone in one ear until it is equally as loud as the tone presented in two ears. This procedure presumes listeners whose hearing is equally good in both ears.

The results of this procedure compare very well with the procedure in which the observer is given a tone in one ear and adjusts another tone until it sounds one-half as loud as the first one. This is the most basic procedure, of course, owing to the nature of what the investigator is trying to do—compare tones by ascertaining their loudness ratio by direct observation.

A third procedure is to find a tone heard monaurally and one that is heard binaurally which have the same value of difference limen. It is assumed that these two tones would have a loudness ratio of 2 to 1. The assumption has been substantiated by results comparable to the other two procedures.

A fourth procedure is possible. It consists in using two tones so different in pitch that they do not activate the same segment of the basilar membrane of the ear. The intensities may be so adjusted as to yield a loudness twice as great as either one alone. Other tones may be equated with the two tones together and with either one alone.

Some laboratories have used one or two of these methods; other laboratories have employed the other methods. The results have permitted those who work in this field to construct a loudness scale based upon the weighted average of the outcome of several studies. In order that this scale might be constructed, certain arbitrary delimiting conditions had to be decided upon.

Since the loudness of a tone of one pitch is a different function of the intensity of the sound stimulus than is the loudness of another pitch, a fixed frequency had to be chosen as a reference. The frequency decided upon was 1,000 cycles. An actual intensity level had also to be chosen for use in setting the unit of loudness. If a low decibel level were to be chosen, the unit would be small; if a high decibel level were used, the unit would be large. The actual intensity level decided upon was 40 db. Thus the unit of loudness was defined as the loudness heard when the pitch of the tone is 1,000 cycles and the intensity level is 40 db. This unit is called the *sone*. Previously, the same author (S. S. Stevens) had proposed the term *phone* as the name of the unit. This term was found to be inappropriate, for the term *phon* has been used in Germany for the equivalent of the decibel. It would be very confusing to have a phone as a subjective unit for loudness and a *phon* as a unit of sound-stimulus intensity.

In general, three different methods have been used in obtaining exploratory data from which the loudness scale is derived. The first of these is the method of *fractionation* and its variations. It is represented by the several procedures described above, in which one tone is presented and a second tone is adjusted until it is sensed as half as loud as the first. The second method is that of *bisection*. With this method the observer is asked to set a variable tone until its loudness is midway between those of two fixed tones. It has been shown that consistency in doing this is possible. Notwithstanding, the results of bisecting do not compare with those of fractionating. Usually the bisections obtained with short intervals agree well with those of fractionation. This cannot be said of results when using long intervals. It seems that observation here involves a paradox. That is, when an observer sets a tone to seem half as loud as another, he is not doing the same thing as setting a tone halfway between a loudness of the given tone and a tone whose loudness is zero. It is obvious that to have the two procedures mean the same thing, the bisection would have to be that of bisecting the interval between the given standard and a hypothetical tone of zero loudness. To make this clear, the diagrams in Fig. 7 are provided.

It has been shown that in bisecting a loudness-interval, the

observer can assume one or the other of two very different attitudes. He can aim at setting the middle tone at a loudness halfway between the other two, or he can aim at setting it so that the ratio of the middle tone to the lowest tone is equal to the ratio of the highest tone to the middle tone, that is, either the arithmetic or the geometric mean can be chosen. With the first attitude, the bisection of the interval between 40 sones and 10 sones would be 25 sones, and with the second attitude it would be 20 sones. It

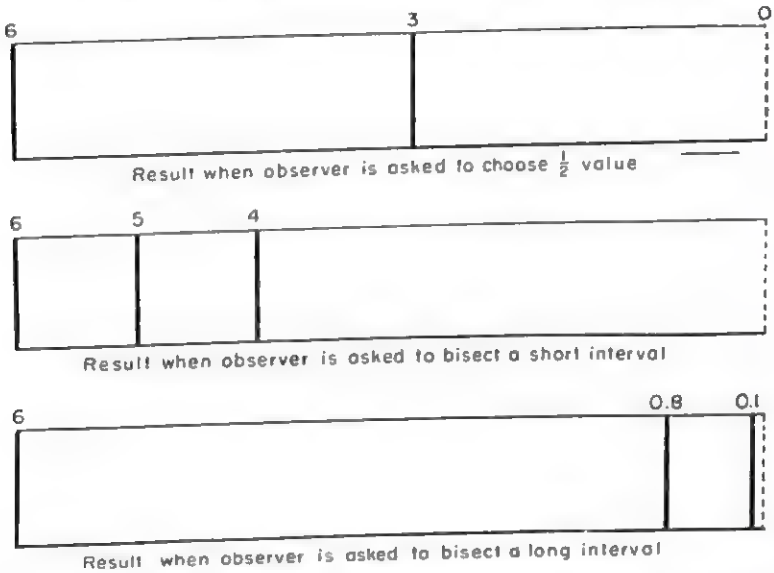


FIG. 7. A schema to indicate the type of response common in judging the value of one-half and in bisecting an interval marked off by two standards. Bisection differs in accordance with whether the interval is great or small. (See text for explanation.)

has been suggested that the inability of observers to hold these two attitudes separate could be the basis for the discrepancy between various sets of experimental results.

The third method is that of judging equidistances. According to this method, an interval in one part of the loudness scale is set so as to be equal to an interval in another part. The intervals may overlap or they may be definitely separated. This is illustrated in one of the diagrams of Fig. 8.

The use of this method indicates that the observers tend to set the loudness of the variable so that equal intervals are actually

marked off. It does not seem that tones are set so as to produce equal ratios (geometric scale). Hence this method tends to confirm the work done by the method of fractionation.

Before the actual unit, the *sone*, was decided upon, Churcher, another worker in hearing, had assigned the number 100 to the loudness of tone of 1,000 cycles and 100 db above the standard reference threshold. We mention this primarily because we wish to introduce an illustration in which a plot of this subjective scale

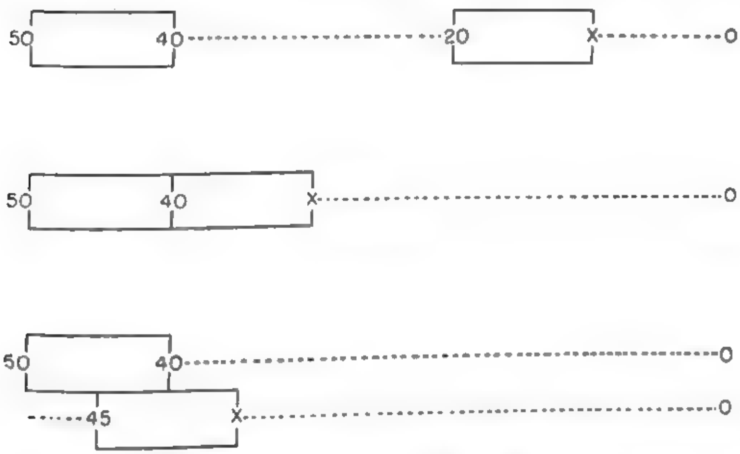


FIG. 8. A schema to indicate three general relations between sense intervals (the separated, the adjacent, and the overlapping). As in the illustrations, the observer is asked to make sense distance 20 to x equal to 50 to 40; or to make 40 to x equal to 50 to 40; or to make 45 to x equal to 50 to 40.

is compared with the decibel scale, and the scale made up of difference limens (DL). It is shown in Fig. 9. In curve *A* is shown the loudness scale of Churcher based upon the unit he chose. The open circles that follow the curve quite well indicate readings which Davis and Stevens obtained when they measured the magnitude of electrical potentials produced in the cochleas (part of the inner ear) of guinea pigs in response to a 1,000-cycle tone of various levels of intensity. The actual values obtained had to be multiplied by a factor to make their value for an intensity of 80 db above threshold equal to that of the loudness curve. When this was done, it was found that the values for all intensities of stimulation, except the very highest, when multiplied by this same factor corresponded very well with the loudness curve.

This seems to indicate that there is in the auditory sense organ a process that is directly recordable and which is approximately the same function of stimulus intensity as the loudness scale itself.

Curve *B* is the decibel scale whose dimensions are fitted to the plot so that it and the loudness scale coincide at a decibel value of

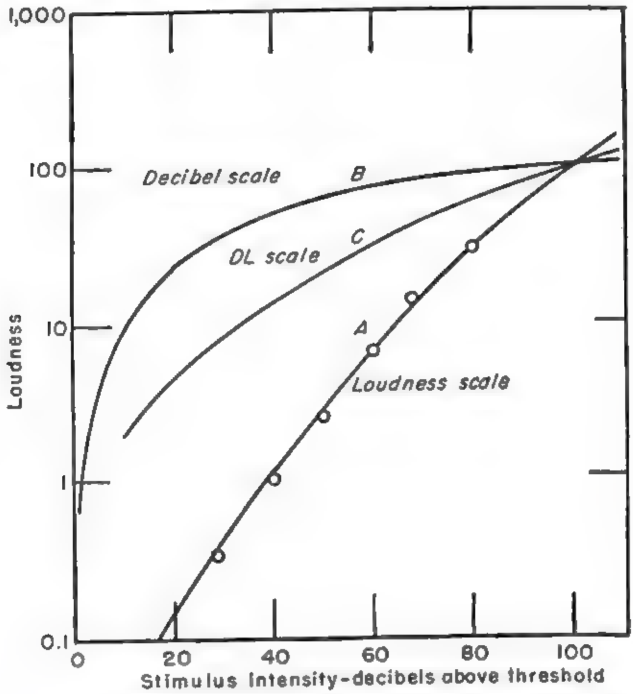


FIG. 9. A comparison between four scales (the decibel scale, the DL—difference limen—scale, the loudness scale, and cochlear response electrically recorded). The decibel scale is shown in *B;* the DL scale in *C;* and the loudness scale and cochlear response in *A*. (See text for discussion.) (*Stevens. By permission of Psychol. Rev., Amer. Psychol. Assn.*)

100. A comparison of the two curves gives the clue to why stimuli stated in decibels do not appear to have the loudness that the number of decibels would indicate were the decibel scale thought of as a subjective loudness scale.

The third curve, *C*, indicates the number of difference limens above threshold as the intensity of the stimulus is increased. The difference-limen function of stimulus intensity is also seen to disagree with the loudness function.

When the logarithm of the cumulated number of difference

limens is plotted against the logarithm of the loudness, curves such as in Fig. 10 are obtained. In this figure, each of the four curves represent findings for a separate pitch. The frequencies represented are 7,000, 4,000, 1,000, and 200 cycles. The points for the 7,000-cycle tone were lifted 0.5 log unit on the

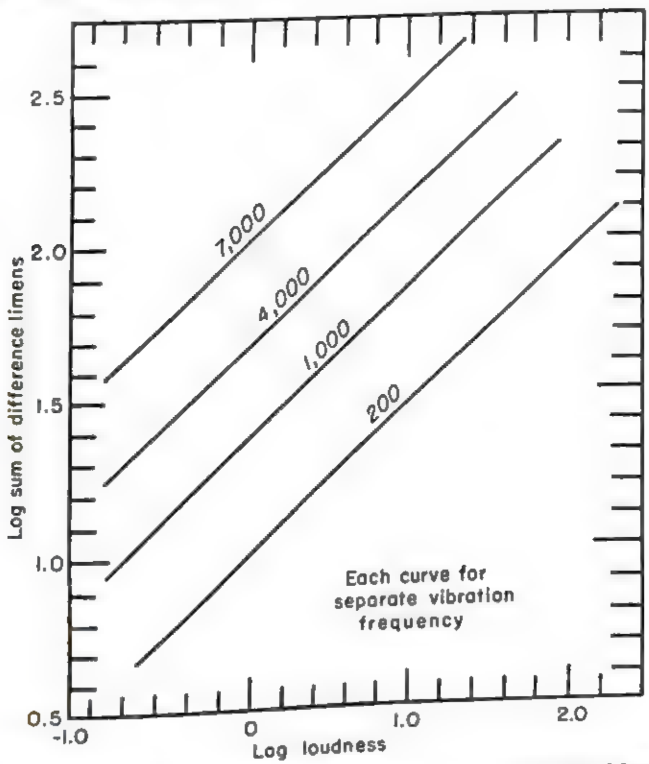


FIG. 10. Graph showing the relation between the logarithm of loudness in sones and the logarithm of the sum of the difference limens. As loudness increases the number of difference limens (DL) per unit of loudness increases. This indicates that the size of the difference limen is not uniform throughout the loudness scale. The curve for 7,000 cycles is moved upward 0.5 log unit for convenience in plotting. The fact that separate curves are necessary for the several frequencies plotted indicates that the size of the DL is not the same for all frequencies. (*Stevens. By permission of Psychol. Rev., Amer. Psychol. Assn.*)

ordinate in order to improve plotting, because the function for 7,000 lies between those for 4,000 and 1,000 cycles.

The relation between the size of the first difference limen above threshold and the frequency of the sound source can be shown in

terms of the number of sones (see Fig. 11). The plot indicates that the difference limens have a very different value depending upon the number of cycles in the stimulus. Tones between 2,000 and 4,000 cycles seem to provide the smallest difference limens. That difference limens are not all of the same value is an important finding, since it was long argued by some that they were identical throughout any given scale.

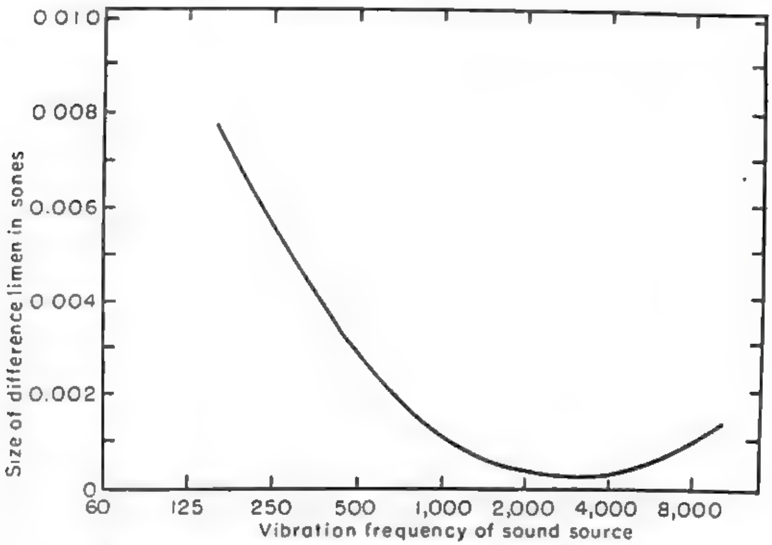


FIG. 11. Graph to show the relation between size of the first difference limen above threshold (in sones) and the frequency of the sound source. It will be seen that the smallest difference limens are to be found for sources with between 2,000 and 4,000 cycles. (*Stevens. By permission of Psychol. Rev., Amer. Psychol. Assn.*)

In Fig. 12, the relation between loudness in sones and the intensity level of the sound stimulus in decibels is shown for each of five different frequencies, 10, 25, 50, 200, and 1,000 cycles. The ordinate representing sones is plotted on a logarithmic scale. If the curves were plotted on an arithmetic scale, they would be of a much different shape and appear as in Fig. 13. Furthermore, with the loudness level covering the great range it does, the whole scale cannot well be plotted arithmetically on a single graph. Note that the range in Fig. 13 covers only from 1 to 100 sones. This comparison between the two forms of plotting is provided so

that you may see for yourself the decided advantage of logarithmic plotting.

This chapter has given you an example of a pioneer work of providing a measuring unit and scale for one attribute of sensation in one sense-modality—hearing. It would appear that this unit and scale have not only a theoretical significance in pure science

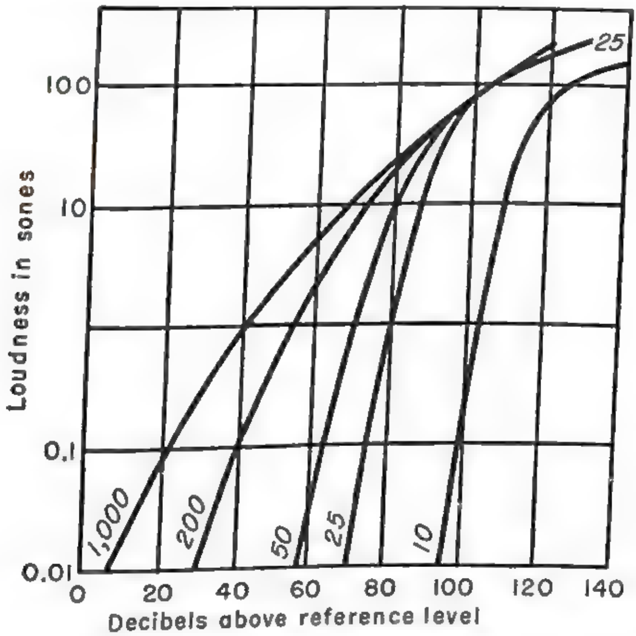


FIG. 12. The relation between loudness in sones and sound intensity in decibels. Loudness is plotted logarithmically. Each curve is for a different frequency. (*Reproduced by permission from Hearing: Its Psychology and Physiology, by S. S. Stevens and H. Davis, 1933,* New York: Wiley.)

but also a practical value. It is presumable that many other attributes of sensation might be handled in the same fashion.

In the next chapter you will be presented with an investigation in which the construction of a pitch scale was achieved. The attribute of brightness in vision might be treated in much the same way as was loudness in hearing. Since brightness involves a number of arbitrary conditions upon which the scale is to be constructed, it might be difficult for some individuals to imagine a use to which a brightness scale could be put. Even though immediate practical use is not foreseen, it would certainly be

inappropriate to object to the attempt to construct such a scale. Every demonstration that sensation can be quantified is, to say the least, of fundamental importance in helping us to gain a better knowledge of the very nature of sensation. It is to be hoped that from now on more and more of this sort of effort will be undertaken.

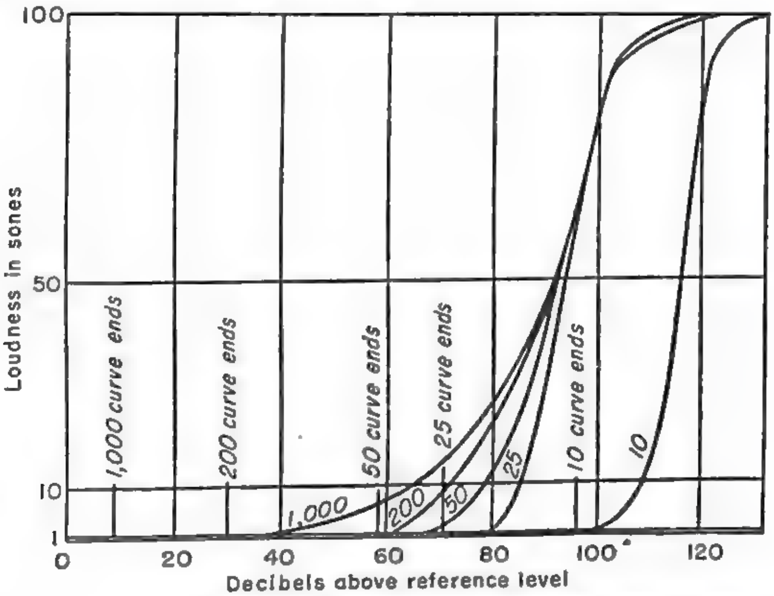


FIG. 13. The relation between loudness in sones and sound intensity in decibels. This is a replot of Fig. 12, in which sones are now plotted arithmetically for comparison.

QUESTIONS

1. Describe the essential nature of the decibel scale.
2. Why are various forms of exponential scales (scales with progressively increasing size of units) used in psychology?
3. Why are the decibel units not suitable for designating experiential properties such as loudness, pitch, heaviness, etc.?
4. What characteristics of the hearing mechanism make possible a variety of ways of comparing the loudness of two tones?
5. Give the specifications necessary in setting up a definition of loudness.
6. Which of the four scales of measurement were used in the present investigation to build the *sone* scale?

7. Contrast sones and j.n.d.'s of loudness of tones.

8. Would the experience of loudness seem to be controlled in the periphery or central nervous system?

9. What is the approximate energy range over which hearing takes place?

10. Give the approximate decibel stimulus values of two common sources of sound.

CHAPTER 9

THE RELATION OF PITCH TO FREQUENCY

Chapter 8 consisted of the description of the way a scale for the loudness attribute of sound was constructed. The description included several methods for studying psychological magnitude that can be applied not only to loudness but also to other aspects of sound and to attributes of other sensations. It was shown that loudness can be quantified on a subjective scale that partakes of all the features of measurement characteristic of physical dimensions.

Workers in the area of sound and hearing have also constructed a scale for pitch. In 1937, Stevens, Volkmann, and Newman related pitch to the stimulus attribute of frequency. The original construction of the pitch scale left certain questions to be answered. Later, Stevens and Volkmann set out to answer them. It is this investigation with which we shall deal in the present chapter.

The Problems. In the construction of the original scale relating pitch to frequency, the following questions remained unanswered: (1) Could the same scale be obtained by different experimental approaches? Would the scale obtained by the method of fractionation compare with the results of an experiment on the bisection of tonal intervals or one on the judgment of equal sense distances? (2) How does the variability among the repeated judgments of a single observer compare with the variability among the average judgments of different observers? To answer this, more observers were needed than were used in the original investigation. (3) Would the results be altered by use of improved instrumentation? The original apparatus was not too satisfactory for tones below a frequency of 120 cycles. Below this frequency, the intensity of the tones fell off rapidly with

decreasing frequency. This led the workers to suspect that the lower end of the scale needed modification. (4) Would an experimental method which directly acquainted the observers with "zero pitch" improve the reliability of the investigation?

Apparatus. To provide working conditions as nearly optimal as possible, the investigators designed what was essentially an electric piano. It contained 20 wooden keys, each 1 in. wide. The pressing of a key produced a pure tone in a loud-speaker. Above each key was a small knob for controlling the frequency of the tone produced. Thus the observers had a repertoire of 20 different tones that could be modified over a wide range by shifts in frequency brought about by turning the associated knobs. The total frequency range of the instrument was 0 to 15,000 cycles per sec.

The instrument was also constructed so that all tones sounded approximately equal in loudness. This was accomplished by an equalizing network contained in the amplifier circuit. The intensity level used was about 55 db.

A 13-in. loud-speaker with a natural period as low as 30 cycles per sec. was suspended at one end of the room. This room was lined with rock wool. The observer sat facing the loud-speaker, 7 ft. away. To augment the tones above 8,000 cycles, the large loud-speaker was assisted by a small crystal loud-speaker suspended 3 ft. in front of the observer. The frequency of the tone produced by pressing any of the keys was read directly from the dial of a wave analyzer connected in parallel with the loud-speakers.

The Pitch Scale Obtained by Equal Sense Distances. Five keys were chosen on the piano. The lowest was tuned to 200 cycles per sec., and the highest to 6,500 cycles per sec. The observer was asked to adjust the three intermediate keys until the pitch distance between each pair of adjacent keys seemed equal. This was actually a task of constructing four equal pitch intervals. As much time as desired was allowed, but no final judgments were to be made until each interval had been compared with every other interval. The observer was to play the intervals both in the ascending and descending orders.

The same procedure was required for two additional ranges of

frequency, namely, 40 to 1,000 cycles, and 3,000 to 12,000 cycles per sec. Ten observers participated in this investigation and made five separate settings of each of the three ranges at different times.

The frequency ranges were so chosen that none of them could be divided in such a way as to represent musical octaves or fifths, the two simplest ratios in musical intervals. This arrangement was meant to remove the temptation from the observers of choosing tones representing these intervals. The investigators pointed out, however, that this precaution was not entirely necessary, for octaves in various parts of the musical scale do not actually sound like equal pitch intervals. It turned out that only infrequently did any observer report a desire to set the frequencies to musical ratios or appear thwarted because he could not do so. In Table 10 is presented the outcome of this procedure.

TABLE 10 *

Observer	40–1,000 cycles			200–6,500 cycles			3,000–12,000 cycles			Avg. V_0
1	174	381	628	865	1850	3260	4090	6780	8440	9.4
2	170	454	718	792	2230	3760	3970	5670	8080	8.6
3	164	327	598	822	2060	3620	4090	5130	7290	6.7
4	167	392	706	751	1696	3270	3820	5260	7540	8.7
5	171	442	718	920	1810	2880	4620	5800	8320	11.9
6	168	391	768	972	1920	3320	4100	5110	7840	5.4
7	142	490	716	848	1859	3120	4410	5700	7480	12.7
8	140	385	732	985	2390	3850	3940	5260	7120	12.2
9	161	404	642	864	2340	3420	3840	5110	7580	13.8
10	154	376	706	846	2065	3430	4210	5440	7740	10.3
Mean........	161	404	693	876	2022	3393	4109	5526	7743	9.9
% variability..	5.9	8.6	6.1	6.4	9.6	6.6	4.4	6.7	4.4	
Mean of means, per cent					6.5					

* After Stevens and Volkmann, *Amer. J. Psychol.*, 1940, **53**, 329–353.

The results shown in the bottom row of the table are the percentage average deviations (V) of the means for the individual

observers about the mean of the group. The mean value of the entire row is 6.5 per cent. This is the *interobserver variability.* The column to the extreme right is, for each observer, the mean of the percentage average deviation (V_0) of his own judgments about his own means. The mean of this column is 9.9 per cent. It is the *intraobserver variability.*

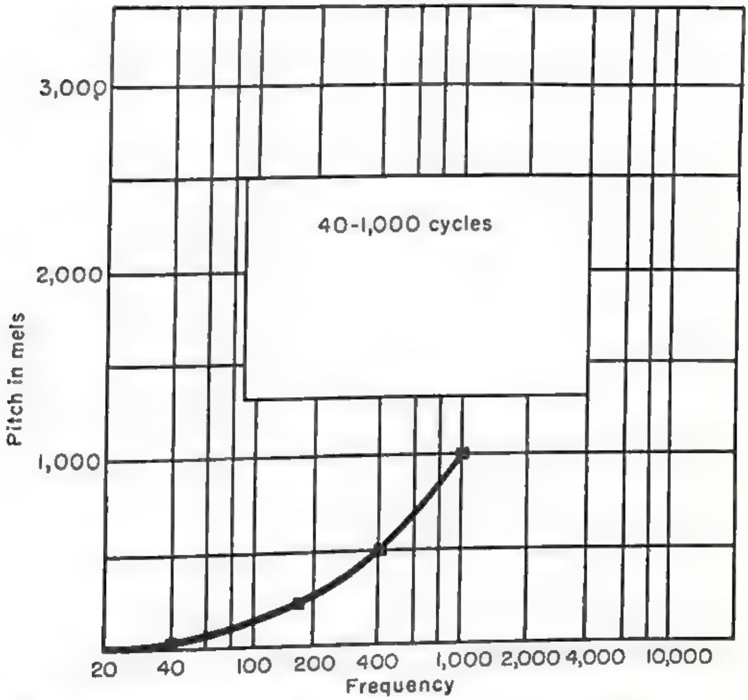


FIG. 14. Relation of pitch in mels to frequency of tones in a range from 40 to 1,000 cycles per sec. This segment was determined by presentation of two tones of fixed frequency (40 and 1,000 cycles) and two other tones which the observer set so as to form three equipitched intervals. This is a portion of the complete curve shown in Fig. 17.

The results are also shown in the series of graphs in Figs. 14 to 17. In Fig. 14, the results of determining equal sense distance (equal pitch intervals) for the range between 40 and 1,000 cycles per sec. are graphed. In the next diagram (Fig. 15), the same is done for the range between 200 and 6,500 cycles per sec. In the third diagram (Fig. 16), the same is done for the range between 3,000 and 12,000 cycles per sec.

It will be noted that these ranges overlap in the frequencies involved. The success in constructing an extensive pitch scale therefore consists in determining whether the overlapping portions of the three ranges display similar values. That is, are the overlapping portions congruent? It will be evident from the diagrams that they are. Figure 17 shows what the complete scale looks like.

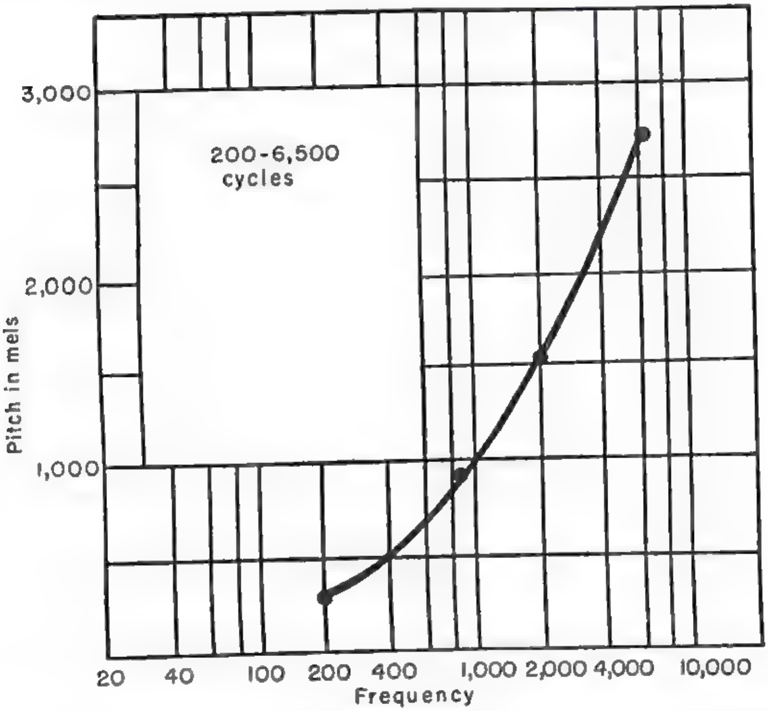


FIG. 15. Relation of pitch in mels to frequency of tones in a range from 200 to 6,500 cycles. Points in the scale were determined as indicated in Fig. 14. This is a portion of the curve shown in Fig. 17.

The ordinates in each of the diagrams (graphs) of the figures just referred to represent *mels*, the units of pitch that the authors, Stevens, Volkmann, and Newman, chose. A mel is 1/1,000 of the pitch heard when a tone is produced by a sound source of 1,000 cycles per sec. and having an intensity of 55 db. The original study employed an intensity of 60 db. Inspection of the graphs, then, will show that it is at a frequency of 1,000 cycles per sec. that the abscissa compares with the ordinate (1,000 mels). This

is, of course, the only place at which the values on the two axes compare. The maximum pitch tested lay at 12,000 cycles. Zero pitch for the scale was obtained by two methods that agreed pretty well. The first method was that of extrapolating downward from 40 cycles, which was the lowest tone used in the scale construction. The second method was that of making a direct judgment of the amount of pitch left in the region below 40 cycles.

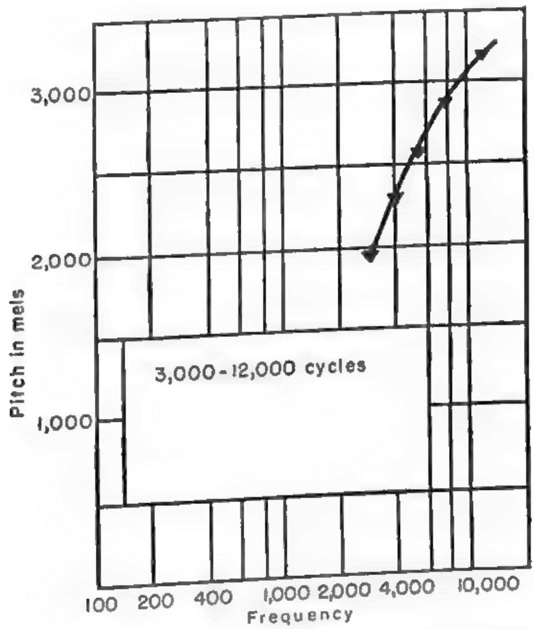


FIG. 16. Relation between pitch in mels to tones in a frequency range from 3,000 to 12,000 cycles. (See Fig. 14 for method of determining points.) This is a portion of the complete curve shown in Fig. 17.

Seven of the observers equated the pitch distance between 25 and 40 cycles to an equal sense distance above 40 cycles. To appear equal, the distance had to extend from 40 to 52 cycles per sec.

The Fractionation of Pitch. The earlier construction of the pitch scale was accomplished by the method of fractionation. The investigators decided to repeat the work on fractionation to see how the results would compare with their findings by the method of equal sense distances (equal-appearing intervals).

The specific procedure used in fractionation in this experiment was as follows: Two tones, each lasting 2 sec., were presented

alternately and were spaced 2 sec. apart. The observer controlled the pitch of the second tone by turning a knob connected to a variable condenser in the beat-frequency oscillator (the instrument producing the tone). Thus, after the first presentation of the first tone (tone actually to be fractionated), if the second tone appeared and was not satisfactorily modified within a few seconds, the observer had additional opportunities to adjust

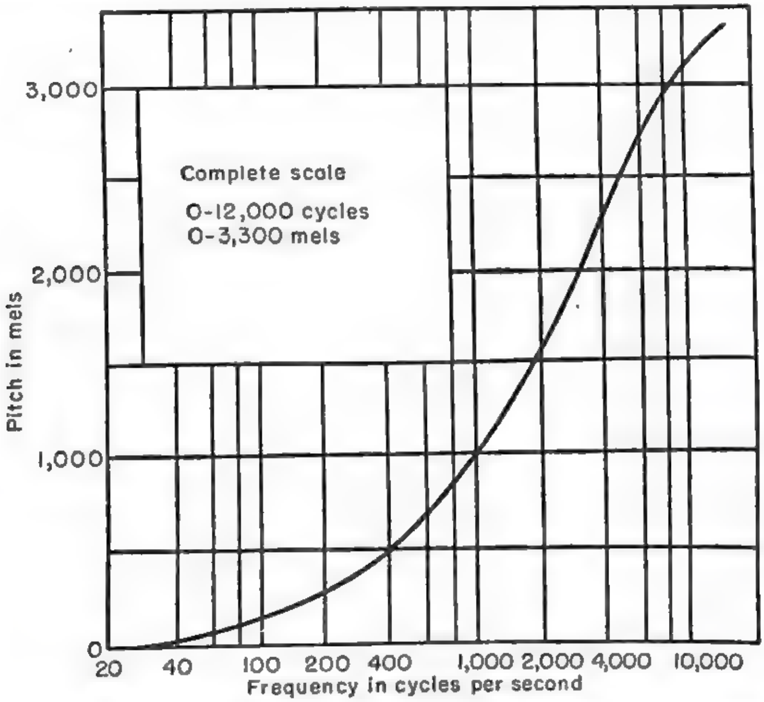


Fig. 17. Relation between pitch in mels and frequency for the complete tonal range used. (*Stevens and Volkmann. By permission of Amer. J. Psychol.*, 1940, 53, 336, *Fig.* 2.)

it so as to appear to have a pitch one-half that of the first tone. The two tones would reappear again and again at regular intervals so that the observer could adjust the pitch of the second one to the point of satisfaction. When the observer felt that he had set the second tone at a pitch one-half that of the first tone, the trial was considered complete. In addition to the procedure just mentioned, the observer could turn a second knob and introduce a tone of 40 cycles per sec. The observer was told that this tone

was very nearly zero pitch. This possibility was included in the procedure owing to the fact that, in fractionation, the zero point is a crucial matter in deciding what one-half of any presented value is. The judgment of "one-half" presumes a zero, and generally zero is a vague affair.

Since actually a 40-cycle tone is not one of zero pitch, it was usable as a reference only with tones of high frequency. The difference between it and actual zero was not great for such comparisons, but the error when using lower tones was naturally relatively great. For example, with a tone of 150 cycles per sec., the observers were not given the chance to use the 40-cycle reference tone.

For this experiment, 12 observers were used, each making five judgments for every reference tone used. The tones used were 150, 250, 500, 1,000, 2,000, 3,000, 5,000, and 10,000 cycles per sec. Four of the five observers used in the initial fractionation experiments of Stevens, Volkmann, and Newman participated in this experiment. The observers agreed that the work was made easier by the use of the 40-cycle tone. The results manifested a greater internal agreement in the present experiment than in the earlier one.

The curve representing the results in the present fractionation experiment paralleled the curve of the earlier fractionation experiment. The fractions chosen in the present experiment were all below those in the earlier one by a roughly fixed amount. This was taken to indicate that the zero pitch involved in the earlier experiment was too high.

The results of the present fractionation experiment agree with those obtained by the method of equal sense distances described in the first part of the chapter.

We have already seen what the variabilities from person-to-person, and in one person from time to time, were in the method of equal sense distances. The average interobserver variability was 6.5 per cent and the intraobserver variability was 9.9 per cent. The investigators found that variability was greater in the method of fractionation. The average interobserver variability was 9.4 per cent and the average intraobserver variability was 15.6 per cent.

The investigators used the pitch scale (mel scale) to predict differential sensitivity in pitch as a function of frequency and concluded that all difference limens (DL) are essentially equal in subjective size. All tones separated by a given number of DL appear equally spaced in pitch. This is in contrast to the findings with the loudness scale. Just noticeable differences (measured by DL) in the loudness scale varied considerably with location in the intensity scale.

The new pitch (mel) scale agrees quite well with the findings of certain other workers (Pratt, Lorenz, Münsterberg), who previously had tested pitch-frequency relationships in local regions of the scale.

QUESTIONS

1. What is meant by "zero pitch," and when does the implication of the existence of a zero pitch become important?
2. How was the problem of zero pitch met in the present investigation?
3. What two kinds of variability were found in the study of the relation of pitch to frequency? Which was the greater?
4. Describe the apparatus used in the present experiment.
5. The use of what caution was mentioned in selecting the frequency ranges for the present experiment?
6. What were the actual frequency ranges used?
7. Contrast the operations of choosing equal sense-distances and of fractionation in the present investigation.
8. Compare the variabilities resulting in the two operations.
9. What is the pitch scale called?
10. How do j.n.d's. for pitch and the experiential units used in the present experiment compare throughout the pitch scale?

CHAPTER 10

LOCALIZATION OF SOUND

Early investigations of sound localization were conducted by use of arrangements generally called *sound cages*. These sound cages consisted of a chair for the observer or subject and a system of tracks about his head. These were circular in form, the

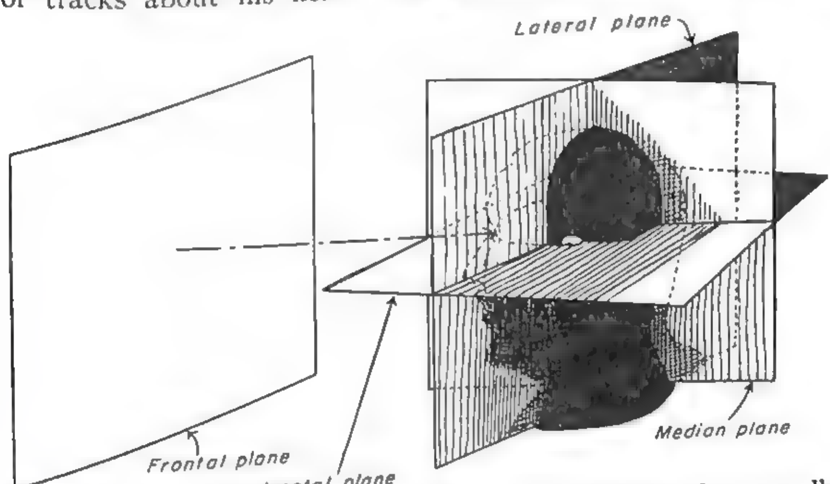


FIG. 18. A schema to indicate the positions of the reference planes usually used in specifying directions from the head. There are three planes that intersect within the head. A fourth plane, the frontal, is used when visual targets are dealt with. It is parallel to the lateral plane and at any chosen distance from the head.

observer's head being at the center of the sphere described by the tracks. Figure 18 indicates the principal planes about the head that are involved by the sound-cage circles. In general, there are three planes, the median, the lateral, and the horizontal. The horizontal plane marks the dividing line between sources that are above and below the ear level. The median plane is the dividing line between sources to the right and to the left of the observer.

The lateral plane marks the dividing line between sources in front of and behind the observer. There are, of course, three *lines* along which these planes intersect. The ends of these lines, however far from or near to the head they may be, mark six *points* at which the planes intersect. The sound-cage circles may be substituted for the edges of the square planes drawn in Fig. 18. Points along these circles are everywhere equidistant from the center of the head.

Sound sources such as mechanical clickers, microphones that emitted clicks or tones, and other similar sources were used at various points along these circles to determine the degrees of accuracy which the observers could attain in judging sound location. Such investigations possessed many limiting features, among which were the lack of electric-generating apparatus for the sounds used and the lack of reverberation control. The experiments were conducted in closed rooms of small dimensions with walls that were not sound absorbent. Despite the limitations, certain facts were established: (1) The individual is able to locate noises better than tones. (2) He is quite able to distinguish right from left in sound origins. (3) He tends to confuse between up and down in the location of sources. (4) He is able to distinguish between front and back.

The excellence of right-left localization suggests the importance of the possession of two ears as receivers at two separate positions. Plugging one ear does, in fact, greatly impair localization. This is paralleled by the fact that individuals with hearing loss in one ear are poor in localization. This impairment seems, of course, to be most marked at first. Later some recovery of ability is usually manifested. Such individuals are subject to confusing right and left, a result very seldom occurring with individuals with binaural hearing.

There is a time difference in the sound reaching the two ears from a lateral direction. This, of course, varies as the direction rises from the zero position (directly in front of the observer) to a 90-deg. position. For a sound source near to the head the equation expressing this difference is $D_t = 0.254 \times 2\,\theta$. D_t is the time difference in milliseconds, and θ is the direction angle, the angular distance from the zero position. For a distant

source, the equation is $D_t = 0.254\ (\theta + \sin\theta)$. The factor 0.254 is 8.75 times the time difference (0.029 msec.) for each centimeter of space difference involved by the two ears, assuming sound travels at 344 m. per sec. The radius of the head rather than its diameter is the space difference involved in this case, hence the factor of 8.75 cm. is used.

The binaural time differences in a sound reaching the two ears is, for example, such as is shown in Table 11.

TABLE 11

Direction angle, degrees	Source near to head, msec.	Remote source, msec.
90	0.799	0.653
75	0.666	0.578
60	0.533	0.486
45	0.400	0.379
30	0.266	0.260
3	0.027	0.027

Figure 19 shows the relation of direction angle indicated by a binaural phase difference of $\frac{1}{20}$ cycle and frequency of sound source in cycles per second. Only at low frequencies is binaural phase difference important in giving directional cues.

The limitations in the experimental equipment of the earlier investigators made the reinvestigation of certain aspects of sound localization in recent years quite desirable. This was undertaken by Stevens and Newman. The description of their work will form the body of the present chapter.

Experimental Conditions. The experiments of Stevens and Newman were conducted entirely in the open air. To avoid, as much as possible, reflecting surfaces, a tall swivel chair was set up on top of a ventilator rising about 9 ft. from the roof of a building on the Harvard campus. Thus there were no vertical reflecting surfaces on any side of the observer at ear level. The nearest horizontal surfaces were about 12 ft. below the observer. The sound source was mounted on the end of a 12-ft. arm attached to the base of the chair. This arm was counterbalanced and could be moved noiselessly in a complete horizontal circle at ear level.

For most of the work a small 4-in. magnetic loud-speaker mounted in a 12-in. square baffle was used to generate the sounds. Only for the lowest frequency (60 cycles) was it necessary to resort to a large Western Electric type 560 cone-speaker to obtain the needed power. This source was mounted at 6 ft. from the observer on the same arm as the first loud-speaker. A

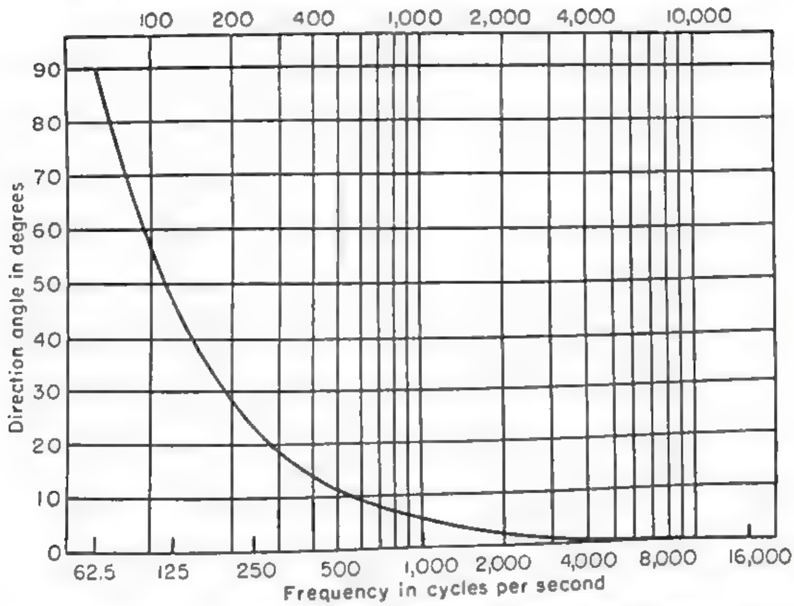


Fig. 19. The relation between the direction angle and frequency of the sound source when two sounds differ in phase by one-twentieth of a cycle. Direction angle is the difference in direction indicated between two tones out of phase by a given amount. This difference in direction varies not only with degree of the difference in phase but with frequency of the sound. Phase difference would be more effective with low tones. (*Data from Woodworth.*)

beat-frequency oscillator, adjustable to the desired frequency, supplied the frequencies to the loud-speakers. The voltage was manipulated by a 7,000-ohm potentiometer. An additional shunting resistance was used to "turn the tone on and off," by regulating its intensity. This procedure avoided the initial and terminal clicks produced by the usual switch.

The results obtained with tones were compared with those obtained with two noises, a hiss and a click. The click was made by briefly applying 45 volts from a battery to the loud-speaker. The

sound heard in this case was a single sharp click with the high-frequency characteristic of the loud-speaker. The hiss was generated by blowing air through a small metal tube, the end of which had been pinched. Part of the energy of the air blowing through this tube gave a high whistle of probably 7,000 cycles per sec., whereas the remainder of the energy was scattered over a wide band of frequencies.

The tones ranging from 400 to 4,000 cycles per sec. had a calculated energy level of 50 to 60 db. The tones in the ranges below and above these frequencies had a calculated energy level of about 30 db.

Procedure. The two investigators acted alternately as observer and experimenter. The tones were presented at predetermined positions in the horizontal plane at ear level. The observer was to name the position of the tone. Thirteen positions on the right side of the observer were used, from directly in front to directly behind the observer. No positions on the left side of the observer were used, since in a preliminary experiment a year before very few reversals from right to left or vice versa occurred. Thus within the limits of time involved, it was preferable to concentrate upon the one side. Ten observations by each observer for each position were made.

The observer attempted to differentiate between the sounds coming from behind and from in front of the strictly lateral position. His success depended upon the sound used. Front and back reversals were frequent, and it was decided that the reversals were not to be counted as errors, for the purposes of the investigation. The *size of the error* was reckoned by taking the difference between the reported position and either the lateral position of the source or the corresponding position in the other quadrant, whichever was the smaller. For example, if the source were directly in front of the observer, the reported position would be considered correct if it were stated as 0 deg., or 180 deg. This way of reckoning is based upon the assumption that dichotic[1]

[1] "Dichotic" is a term indicating that which affects the two ears differently, as in the conveyance of one sound to one ear and a different sound to the other at the same time. It is a term that in audition parallels in many ways the term "binocular" in vision.

differences of phase or intensity provide a basis for general lateral localization.

Localization as a Function of Frequency. The mean of the errors made by the two observers at each frequency is shown in Fig. 20. It will be noted that the errors of localization made at the low frequencies are comparatively constant. The errors

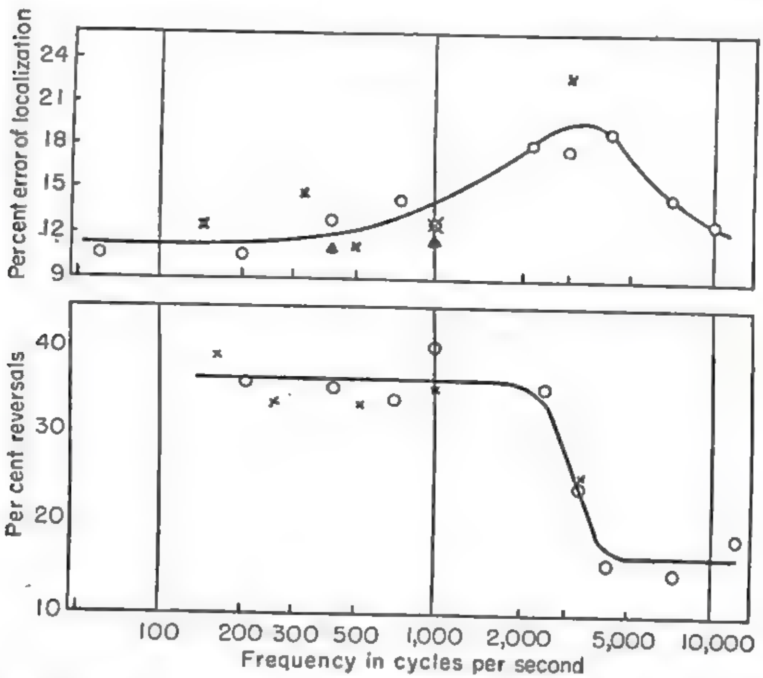


FIG. 20. (Top) The relation between error of localization and frequency of sound source. The maximum percentage error occurs with tones in the neighborhood of 3,000 cycles per sec. (Bottom) Relation between per cent of reversals in front and back sound sources in relation to frequency of sound sources. Reversals are great (about a chance value) below about 2,000 cycles. The reversals are only one-third the number predicated on a chance basis for tones above 4,000 cycles. (*Stevens and Newman. By permission of Amer. J. Psychol.*, 1936, 48, 302, *Figs.* 2, 4.)

become larger as the 3,000 cycles are reached. Above 4,000 cycles localization begins to improve again and is almost as accurate at 10,000 cycles as at 1,000 cycles.

The existence of a maximum for errors in localization in the middle range of sound frequencies calls for some explanation. The investigators were able to supply this from the work that had

been done previously by others. Taking into account the size and shape of the head, certain intensity and phase differences in the sounds reaching the two ears are to be expected. From both theoretical calculations and from the literature on the subject, phase differences are found to be effective as cues with sounds of low frequency. The limit of effectiveness in phase difference apparently reaches its peak at a frequency in the neighborhood of 1,520 cycles per sec. Steinberg and Snow showed that the difference in intensity of sound energy reaching the two ears tends to increase with frequency. This differential begins somewhere below 1,000 cycles per sec. and begins to progress rapidly beyond 3,000 cycles per sec. It would seem then that at 3,000 cycles per sec., phase difference is too little to help greatly as a cue to localization and the differential in intensity has not yet risen high enough to be very effective. This being the case, the region of poorest localization would manifest its peak at or near 3,000 cycles per sec., owing to the relative absence of both of the bases for cues.

Localization as a Function of Position. Since the 1890's it has been recognized that the localization of sounds directly to the front or behind the observer are localized most accurately.

In the present investigation, when the frequency of reversals was taken as a measure, the ability of the observers to distinguish between front and back was striking. The relation of the percentage of reversals and the frequency of the sound source is shown in Fig. 20. It is noticeable that the frequency range is divided into two distant regions separated by a narrow transition. The mid-point of this transition is at about 3,000 cycles per sec. For tones below 2,000 cycles, the range in which localization is based on phase differences, distinction between front and back is only a little above chance. On the upper side of the transition (4,000 cycles and higher), the percentage of reversals is but one-third of those expected on a chance basis. This is confirmation of the idea of the two-mechanism basis for localization just mentioned (phase difference and intensity difference). The ability to distinguish between "in front of" and "behind" with high-frequency tones would seem to be largely a function of intensity.

Several checks on this hypothesis were made. One of these involved using a continuous high tone. When this was swung around the observer, it appeared to be less intense when behind him. In another case, intensity of the tone was varied from trial to trial. As a result the percentage of reversals for tones in the 3,000 to 7,000-cycles-per-sec. range increased from 18.6 per cent (in constant intensity procedure) to 47 per cent in the present variable intensity procedure. This would suggest that the observer had formed a subjective standard of intensity in the main experiment after a very few trials. Subsequently the tones in back of him seemed relatively weak, and those in front seemed relatively strong. The basis for the difference might be expected to lie in the shape of the outer ear. Its projection would tend to produce a sound-shadow and thus reduce the energy reaching the inner ear from behind the head.

Localization as a Function of Sound Complexity. The series of trials in which the hiss and the click were employed confirm the commonly held conclusion that sounds of complex wave composition are more easily located than pure tones. The localization of the hiss was especially definite. The experimenters state that it was almost as definite as though one were looking at the source of the sound. The average error for the hiss was only 5.6 deg. It was 8.0 deg. for the click. The errors of localization rose to as high as 16.0 deg. for the pure tones.

The fact that the localization of the click is better than that of the pure tone demonstrates the necessity of the originally mentioned technical precaution of eliminating the usual click at the onset and termination of the pure tone. Without this absence of click, no real indication of the accuracy of pure-tone localization could have been achieved.

When a check was made to answer the question of whether there were any intensive and qualitative differences in the hiss when generated in front or behind the observer, the results were positive. The hiss was louder when coming from in front of the observer. It also sounded more like *shh* from in front, and like *sss* from behind. Neither this qualitative difference nor the difference in intensity came to attention during the main series of judgments. The sound-shadow of the head undoubtedly had a major

part in the qualitative differential in the hiss from behind and in front of the observer.

QUESTIONS

1. Describe a sound cage.
2. What are the planes and meridians often used in dealing with hearing?
3. In what directions are most accurate sound localizations expected?
4. What improvements in procedure and other experimental conditions were introduced in this experiment, compared with those in the past?
5. What kinds of stimuli were used?
6. Which type of stimulus was most accurately localized?
7. What differences in experimental results between stimuli used in front of and behind the observer were discovered?
8. What other important variable aside from direction was involved in the sound-localization investigation?
9. Are all tones equally well localized?
10. What is a sound-shadow?

CHAPTER 11

THE CONTROL OF LOCATION AND PERSPECTIVE IN SOUND

In the foregoing chapter some of the facts and experiments concerning the localization of sound were presented. The investigation cited had to do simply with the detection of the direction of sound sources in a natural, uncomplicated situation. There are other aspects to localization that come into play when today's synthetic imitations of "natural" situations are involved. The problems pertain to what has to be done to make the synthetic or otherwise "manufactured" situations resemble the type which we call natural. Among others, there are problems pertaining to the movement that artificial sound sources must make to copy the movement of seen objects that are supposed to be emitting the sound. In these days a great many sounds are instrumentally recorded for subsequent reproduction. This procedure involves certain integrations of knowledge in acoustical engineering and in the nature of hearing. Since engineering principles and information loom large in such cases, their description may seem to belong to treatises on engineering rather than in the present textbook. This is not necessarily the case, for these situations provide manipulative conditions for understanding the processes of human hearing that are found nowhere else. Such situations serve also in a very graphic and convincing way to show how close psychology as a science and the constructive pursuits in everyday life are knit together. Such studies exemplify one of the aspects of industrial psychology that is neglected in the present concentration of attention upon personnel selection. On this account, the present chapter could well serve in this text as one representing the field of industrial psychology. We shall leave that task, however, for Chap. 48.

One example in the area to which we are alluding is the stereophonic effect. In certain natural situations, such as in listening to a symphony orchestra, the listener senses the spatial relation of the various instruments. This is a type of stereoeffect. It pertains not only to positioning objects both as to right and left from the listener, but also as to relative distance. The listener thus not only sees an orchestra occupying three-dimensional space, but he experiences three-dimensional sound effects. Auditory as well as visual perspective is involved.

When music is not to be heard directly but is recorded for later reproduction, the aim is to preserve the dimensional effects of the complex sound situation. When but a single microphone is used to pick up the music, the possibilities of doing this are scant at best.

One of the ways for recording sound in auditory perspective is to use a form of binaural pickup. Sound, in such cases, is picked up by two microphones placed in the positions of the ears on a manikin. If two earphones are used by a listener in some other room, one earphone connected to each microphone, the observer will receive the necessary stimuli to hear sounds in perspective. For example, if someone walks around the dummy as he talks, the listener in the distant room will be given the compelling effect of a speaker walking around him. This binaural pickup is certainly not simple enough to be practical where sounds are recorded for subsequent playoff on records, etc. Avoidance of two channels, two records, etc. (one for each ear) is certainly demanded.

In practice, two or three microphones—loud-speaker channels—are used to provide fair auditory perspective. In one investigation, Steinberg and Snow set up a three-microphone, three-loud-speaker system to determine its effectiveness in communicating the physical position of an unseen performer (caller) to a group of 12 listeners 50 to 60 ft. beyond a gauze screen in an auditorium. The caller, at random, assumed each of nine positions in a square formed by three positions in each of three rows in a pickup room. The sound of the caller's voice was transmitted to three microphones spaced parallel to one side of the square and parallel to the screen that was out on the stage.

Behind the screen, and parallel to it, were three loud-speakers. These were connected to the microphones in each of five ways. By throwing switches the connections could be shifted from one combination or pattern of connection to another: (1) The simple form of connection was that of each microphone being connected to its own loud-speaker. (2) A reduction of this was tried when only the two end microphones and the two end loud-speakers were used. (3) Another combination involved two microphones and three loud-speakers. In this case, the three loud-speakers were connected together and the two microphones were introduced into the circuit, one on either side of the middle loud-speaker. (4) In another case only two loud-speakers were used with all three of the microphones. The three microphones were connected together, and the two loud-speakers were connected in, one on either side of the middle microphone. (5) In a final change of circuit a more elaborate system of introducing the relative energies from the three microphones was tried. All three microphones were connected together. All three loud-speakers also were connected together. Connections from the microphones to the loud-speakers were made from points on either side of the middle microphone to points on either side of the middle loud-speaker.

The object was to distribute the energies in different relative amounts from the three microphones to the three loud-speakers, from two microphones to the three loud-speakers, or from three microphones to two widely separated loud-speakers. One lateral loud-speaker was a little more than 20 ft. from the middle one; the other lateral loud-speaker was about 22.5 ft. from the middle one. The width of the auditorium was about 60 ft.

The lateral distribution of the caller's voice from one position to another of the nine was communicated quite well in every case, although not equally so. The front-to-back positions were differentiated best by hookup 5 and poorest by hookup 2.

The lateral localization was derived from the differences in the binaural intensities produced in each case by the hookup used. The front-to-back localization was supposedly derived from another set of factors. It has been found that with a single-channel system, an increase in the ratio of the sound reaching the

microphone directly to the amount reaching it by reflection from the walls causes the sound source to appear nearer to the listener. This is to say, relatively more reverberant sound is heard in a room when the source is farther away.

Maxfield's Investigation. The research that we have in mind for more special consideration in this chapter is that of Maxfield, in which the factors of sound localization are matched with appropriate visual factors involved in photography. Not only certain factors which we have already mentioned were involved, but also visual perspective as it is produced by camera distances, focal lengths of lenses used, lighting effects, and sound-absorbent properties of the sets used.

Purpose. The aim of Maxfield's experiments was to determine in principle what manipulations of lenses, camera distances, and microphone positions were necessary to provide realism in talking motion pictures. The particular effect sought was to have the sound seem to come from where the source in the picture demanded it. We prefer not to call the attempt that of seeking to produce an illusion, but to speak of the matter as one of trying to determine what the necessary and sufficient conditions are for given desired effects. In the scientific sense, all effects are *real*. None are illusory. Since both the picture and the sound convey perspective, it is necessary to obtain correspondence between the distance of the object in the picture and the distance or position at which the sound appears to originate. Two types of motion are usually brought into play on stage sets. They are forward and backward (receding and proceeding) motion, and motion from one side of the picture to the other. The sound and the object in the picture must appear to move together.

Setups and Conditions. During the making of several pictures with which Maxfield was connected, the following data were collected: (1) focal length of camera used; (2) microphone positions; (3) construction materials of set walls; (4) conditions such as open doors, nature of sound stage, etc.; and (5) judgments of engineers and art critics regarding the effects when the pictures and sounds were reproduced in the reviewing room. In some instances several cameras were used. In such cases it was

possible to select which of the two or three pictures best matched the sound. The camera focal lengths and the relative positions of the microphones were plotted against each other. When this was done, there was a trend indicating that the longer the focal length of the lens, the nearer the microphone must be placed to the speaker. There was considerable spread in the points on the plot, however.

In the attempt to discover the reason for the spread, the points for sets possessing walls of different materials were replotted using separate symbols for sets with highly sound-absorbent material (class 1), average absorbent material (class 2), and relatively nonabsorbent material (class 3). When this was done, it was immediately obvious that the points representing live sets (class 1) were, in general, lower on the plot than those for the normal or average sets. The judgment of whether the sets were live, normal, or dead were made by ear by Maxfield.

Later, Maxfield made measurements of the sound-absorbing properties of materials from which sets are usually constructed. There, coefficients of absorption varied widely. The materials having coefficients of medium value corresponded well with the materials of sets previously judged average. Likewise, the materials having absorption coefficients of low value matched well with the materials of the sets previously judged live.

Figure 21 is a formalized plot of the relations between focal length of lens and microphone position for sets of the three classes (1, 2, and 3). The middle line represents microphone positions at the camera whose lens has a focal length of 35 mm. when the set has an average sound-absorption coefficient (an average set). The abscissa represents lenses of various focal lengths, and the ordinate indicates microphone positions in per cent of distance from object to camera. For example, if the microphone should be placed approximately at the camera having a 35-mm. lens, the microphone should be placed 42.5 per cent of the distance from the object to a camera having an 80-mm. lens, and 34 per cent of the distance if the camera had a 100-mm. lens.

If the materials of the set walls are live, the percentage of the distance from the object to the camera at which the microphone

must be placed must be reduced. For an 80-mm. camera lens the microphone would have to be placed at 30 per cent of the distance. On the other hand, for a live set (highly reverberant), a 25-mm. lens can be substituted for a 35-mm. lens and percentage distances used, if other lenses of greater focal length are later substituted. The reverse is true for dead sets. For them

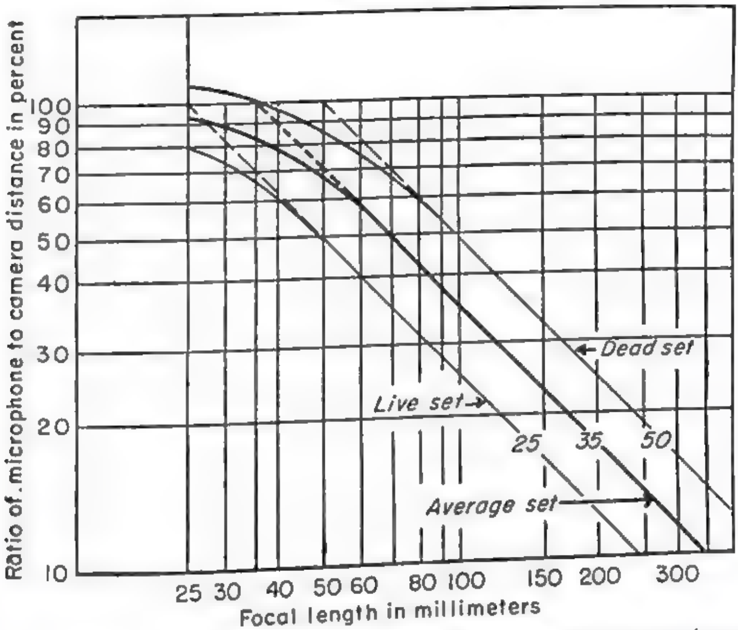


FIG. 21. Graphs to show the distances from a photographed object (speaker, etc.) that a microphone should be placed, depending upon the focal length of the camera lens. The center curve is for a lens whose focal length is 35 mm. If the pickup is appropriate for this lens when the microphone is at camera distance, the microphone must be placed at other distances from the sound source for lenses of other focal lengths. If the focal length is greater than 35 mm., this distance is some percentage of the camera distance below 100. The other curves are for sets (stages or scenes) where the sound absorption is either greater or less than average. For these, 25-mm. and 50-mm. lenses may be used instead of a 35-mm. as references. (*After Maxfield. J. acoust. Soc. Amer.*)

a 50-mm. lens may be substituted and the upper of the three curves in Fig. 21 consulted for microphone positions when other lenses are substituted.

In case the function of focal length of camera lens might not be apparent, Fig. 22 is presented. A short focal-length lens

covers a wider scene than a longer focal-length lens. This is obvious from comparing diagrams A and B. In each of the diagrams there is an object shown in two positions—far and near. As the object moves from far to near in situation A, its image increases from the one size to the other as indicated by the lines at the bottom of the diagram. The same thing, of course, happens in situation B, but the ratio of change in image size is less.

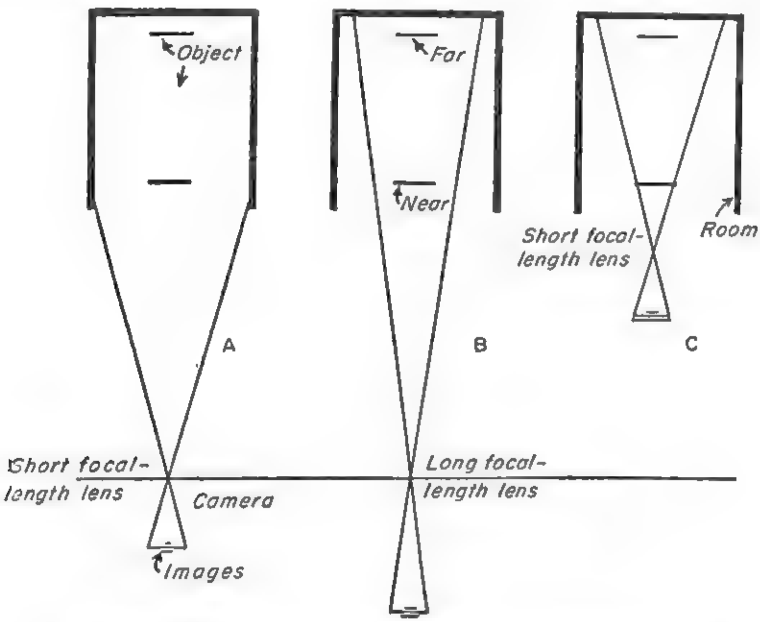


FIG. 22. Diagram to indicate certain properties of lenses of short and long focal length. (See text for discussion.)

In B, the perspective effect is relatively less. With a short focal-length lens and a long focal-length lens at the same distance from a scene, the latter will make the scene seem closer. It is for this reason that microphone positions have to be adjusted according to actual distances from photographed objects and in accordance with focal lengths of lenses used. When a sound is right for a picture taken with a lens of given focal length, it must be picked up at a different (lesser) distance for a picture taken with a lens of greater focal length.

Figure 23 represents an actual set with walls of average sound absorption. In diagram A, the position of the long shot is shown.

In diagram B of the same set the arrangement for close-ups are shown. In *A*, the action is shown so as to indicate the motion and thus the pictorial depth involved. In position 1, two individuals carried on a dialogue. One of the individuals walked from position 1 to position 2, in the rear corner of the room, to talk with another person. Soon afterward the person who came from position 1 left the scene through a door as indicated. The

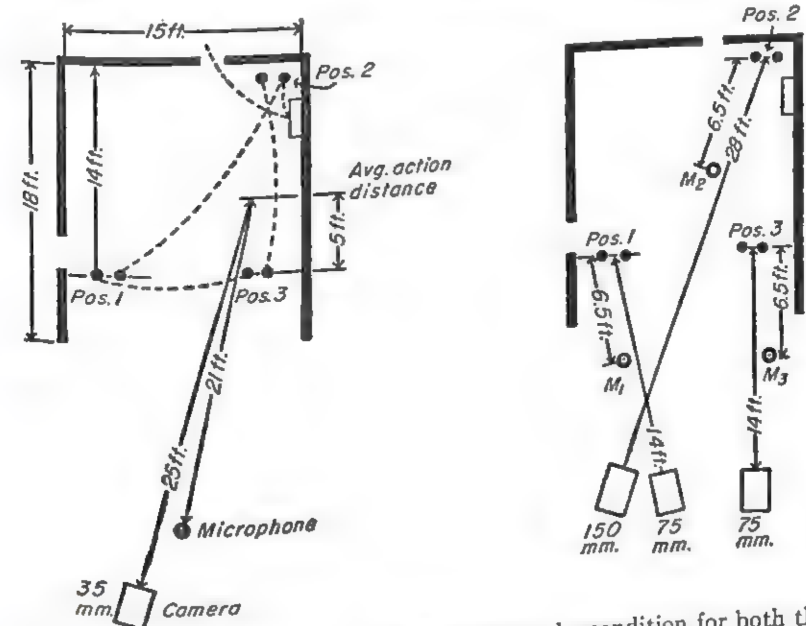


FIG. 23. Two diagrams of a single set, showing the condition for both the long shot and the close-ups. The important factor in each case is the relative distances of the microphones from the object to camera. The additional factor is the sound-absorbent properties of the walls of the set. (See text for fuller discussion.) (*Maxfield. J. acoust. Soc. Amer.*)

other person walked forward to talk with the second person, originally in position 1, who came across the set from left to right to position 3.

There was only one microphone in action at any one time for the close-ups. The microphone for the long shot was in action during the motion from positions 1 to 2 and from 2 to 3. The close-up for position 1 used the microphone at M_1. The close-up for position 2 used the microphone at M_2, and the final close-up at position 3 used the microphone at M_3.

Editing the film and sound track involved placing the close-ups in their proper places within the long-shot sequence. The realism of the results was quite satisfying. In one of the initial attempts at editing, the long-shot sound track was paired by mistake with the close-up for position 2. In other words, the sound was such as recorded by the microphone at M, when it should have been the pickup made by the microphone at M_2. The lack of realism was disconcertingly obvious. In this case the voices of the speakers seemed to come from outside the room through a window behind them. No observers of this mistake were able to hear the voice as coming from the position of the speaker, even when the loudness of the voice was artificially (instrumentally) increased greatly.

Maxfield found that the effect of speech being too close, although obvious, does not seem to be quite so disturbing as to have it too far away. Destroying the realism by having the sound come from too close makes the sound appear to come from the loud-speaker horn just behind the motion-picture screen. That is, it comes from the actual position of the source producing the sound, when the scene is shown on the screen.

It will be noted that the microphone for the long shot was not actually at the camera position but somewhat closer to the actors (sound sources). This is in accordance with practice and is illustrated by the curved positions of the graphs in Fig. 21. Moving the microphone up toward the scene a little was found advantageous. The reason Maxfield gave for this was the avoidance of camera noises and the reflectance of sound from the walls of the enclosing sound stage. What would be slight discrepancies at short filming distances are not detectable in long shots.

Figure 24 presents another set in which the actor enters a side door at the far end of the room and advances to position 1, then to positions 2 and 3. The actor is a girl, singing as she advances. It will be noted that although there is but a single camera position, three microphone positions are used. These positions do not vary as greatly in absolute distance as do the singer's positions (1, 2, 3). But each of the microphone positions are in the neighborhood of 47 to 50 per cent of the distance from

singer to camera. This is in accordance with the plot in Fig. 21, which indicates a percentage distance of about 47.

This investigation, although not conducted in so formalized a fashion as the usual scientific laboratory experiment, was actually

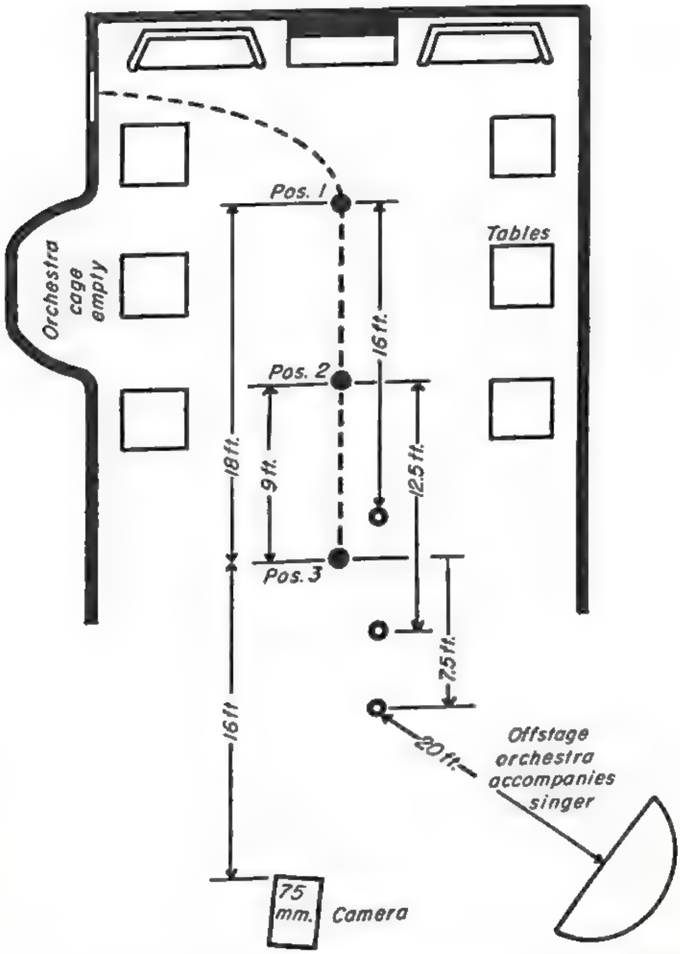


FIG. 24. Diagram of a stage set to indicate microphone positions for each of three actor positions in relation to distance from actor to camera. (*Maxfield. J. acoust. Soc. Amer.*)

experimental in nature. It ought to be suggestive of one kind of problem met with in industrial situations and also lead to your conceiving other problems in the area of acoustics and audition.

QUESTIONS

1. What is the stereophonic effect?
2. How is some degree of this effect produced in the recording of sound?
3. What did the experiment of Steinberg and Snow show?
4. With a single channel pickup, what factor influences the characteristic of the sound heard?
5. What factors did Maxfield manipulate in his investigation on the matching of sound and motion pictures?
6. State in other words what Maxfield was trying to do in studying the matching of sound and picture.
7. What essential difference is there in a scene photographed with a long focal-length and a short focal-length lens?
8. What is the connection between the resilience of the walls of the room ("movie set") and the focal length of the lens used?
9. What effect is produced by having the microphone too far from the object (sound source) toward the camera?
10. What effect is produced when the microphone is too close to the sound source?

CHAPTER 12

BINOCULAR FACTORS IN SPACE PERCEPTION

It has long been known from optical considerations that as the two eyes move to and fro, the retinal images of stationary objects shift their positions and vary in size and shape. Even among those who have been aware of this and have had some interest in visual perception, few have examined the consequences in minute detail. A careful examination is necessary, however, if one wishes to understand visual space perception. If one considers the relations between retinal images and spatial experiences, he will discover a number of lawful relations. These changes in the sizes, locations, and properties of the retinal images constitute the stimulus factor in the observer's experience of a visuospatial world in which objects change size, position, and orientation. That binocular spatial perception becomes abnormal when the relations of retinal images on the two retina are altered by lenses or by prisms has long been known. Not a great deal could be done to study the effects of abnormal retinal images, because when images were intentionally distorted by lenses, the images tended to preclude stereoscopic interpretation by the observer. Even when such interpretation does result, the blurring makes it impossible to determine *accurate* relations between a given interpretation and the images involved. The same difficulties arose in the use of prisms.

Nowadays we have at our command another kind of lens, called the size-lens, that does not blur images, even though it can be used to distort them. It is the purpose of this chapter to introduce you to its nature and effects. A review of the main features of correlation between object localization and retinal images is to be given here to preface the specific experimental work which we wish to consider more in detail. This review will be given by a series of diagrams.

Lenses. All lenses may be classified into two groups—power lenses and size-lenses. The lenses with which you are familiar are power lenses. These magnify or diminish the apparent size of objects and they converge or diverge rays of light. That is,

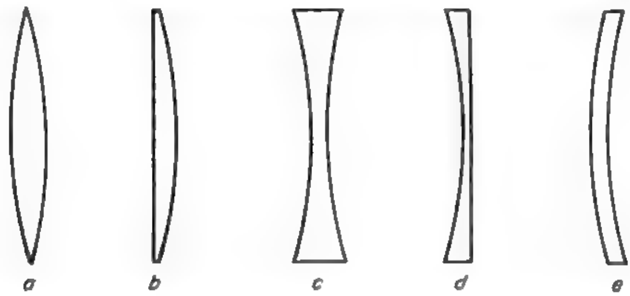


FIG. 27. Power lenses shown in *a*, *b*, *c*, and *d*. Size-lens in *c*.

they can be used to *focus* and are described by their focal lengths, etc. Figure 27 illustrates the common forms of power lenses—the biconvex, the planoconvex, the biconcave, and the planoconcave. This same figure shows what is meant by a size-lens.

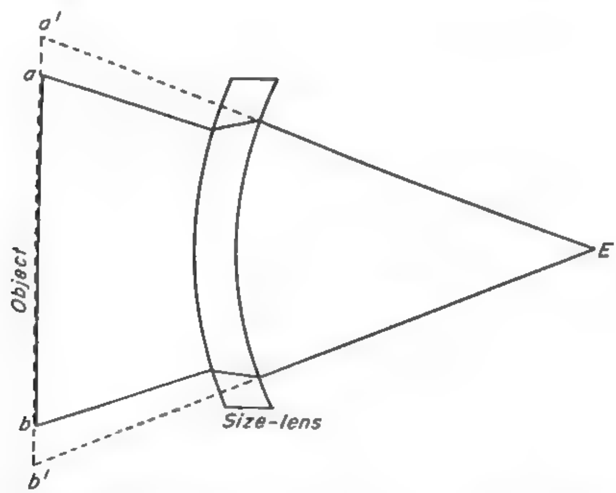


FIG. 28. Diagram to show effect of size-lens. Light from points *a* and *b* reaches *E* but appears to come from points *a′* and *b′*. This produces a greater apparent over-all size of *ab*.

Size-lenses are formed by bending or cupping material, the surfaces of which are parallel to each other ("plano"). If a piece of plano material, such as a piece of window glass, were

curved in one dimension, it would look like *e* of Fig. 27, and its optical properties would be as indicated in Fig. 28. From this, it is obvious that the significant feature of such lenses is their magnification. If the material from which they are made is bent along one axis, the lenses are called *meridional* size-lenses. If the material is cupped, the lenses are called *over-all* size-lenses.

Size-lenses and Their Effects. What the effects of size-lenses are constitutes our problem in this chapter. To solve this, we ask ourselves, of course, something about their optical properties.

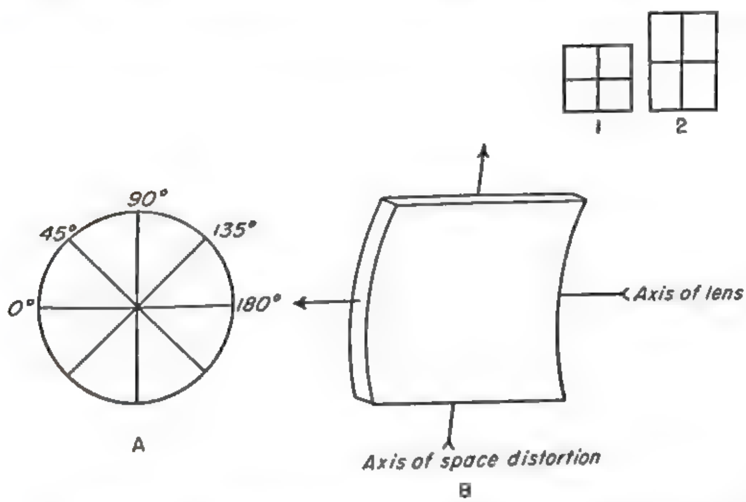


FIG. 29. Diagram A indicates the manner of labeling lens axis in terms of degrees. B indicates the relation of axis of the bending of material to form a size-lens, and the axis of optical distortion produced. Inserts 1 and 2 show the elongation produced in the image of an object, along the axis of distortion.

That is, how do they treat light? Furthermore, we ask in how many ways they can be utilized to modify visuospatial experiences. When meridional size-lenses are worn on both eyes, it would be supposed that they should be rotated so as not to produce the same optical effects in both eyes.

With meridional size-lenses several facts must be kept in mind in order to relate *lens axis*, *distortion of retinal image*, and the *visual effect that results*. Lens axis is measured in degrees that are counted from the left-horizontal meridian of a circle, as indicated in diagram A of Fig. 29. Since the axis of a lens is a diameter,

enumeration need not extend beyond 180 deg. The axis of the size-lens shown in diagram B could be called either 0 deg. or 180 deg. By custom it is called 180 deg. Thus, although the axis of bending of the lens is 180 deg., the effects lie in an axis at right angles to it, or in the vertical axis. This vertical distortion of the image of a square is shown in inserts 1 and 2. (For graphic purposes the square is bisected vertically and horizontally by lines.)

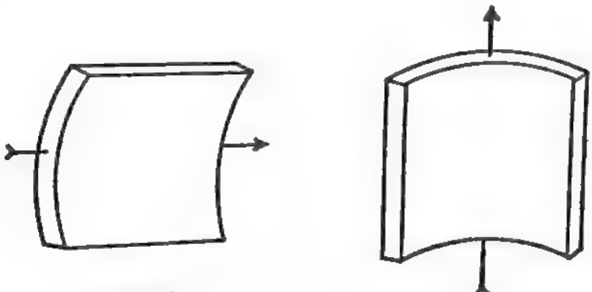


Fig. 30.

Figure 30 shows two principal axes for size-lens use. Figure 31 illustrates what happens when a meridional size-lens (axis 90 deg.) is applied to one eye (right eye). The size-lens magnifies the image in it. The observer sees the object as single. The position of the ends of the object are not at AB as when no size-lens is worn, but at $A'B'$ where the "visual" lines projected outward from the eyes intersect. *As a rule, an object looks to be where it would be were the light to come straight from it to the eyes.* Thus the lines used for the right eye are the projections of the rays of light after they have passed through the size-lens. The intersection points of the projections of the lines for the two eyes are at $A'B'$. These points represent the ends of a real object that would be in the position of the representation $A'B'$. Thus it is that the wearing of a size-lens (axis 90 deg.) before the one eye tends to make objects look farther away on the side on which the lens is worn. Such objects also look larger. Walls, floors, ceilings, etc., are subjectively tilted or rotated out of their usually seen positions in accordance with this rule.

If meridional size-lenses are worn before both eyes, but with

their axes in the oblique positions, still other effects will result. If the axes of the lenses are turned toward each other at the top, the floor will seem to tip upward from where one stands. If the axes of the two lenses are turned outward at the top (this provides magnification in the axes of the eyes that tilt *toward* each other at the top), the floor will appear to tip downward or recede. Figure 32 shows two lenses whose axes tip away from each other

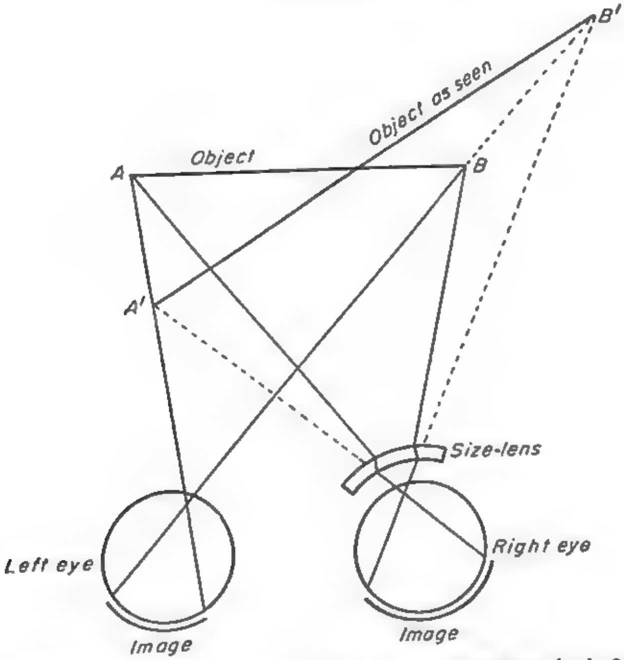


FIG. 31. Diagram to show the effect of wearing a size-lens (axis 90 deg.) in front of the right eye. Note that the object appears where it would be if the light came straight from it to the eyes, as judged from the direction light enters the eyes.

at the top. It also indicates the distortions produced in the retinal images of a twice-bisected square.

The Leaf-room. The set of conditions under which to observe the effects of wearing size-lenses is provided by what has come to be known as a *leaf-room*. It is a 7-ft. cubicle lined with artificial (theatrical) leaves. One of the four sides of this room is left open. Midway between the two side walls, as shown in Fig. 33, the observer takes up his position. If he is given a meridional size-lens (axis 90 deg.) to wear in front of the right eye, the room

will appear distorted as indicated by the right-hand drawing. The back wall of the room appears to be farther away at the right than at the left. The ceiling of the room appears to be higher

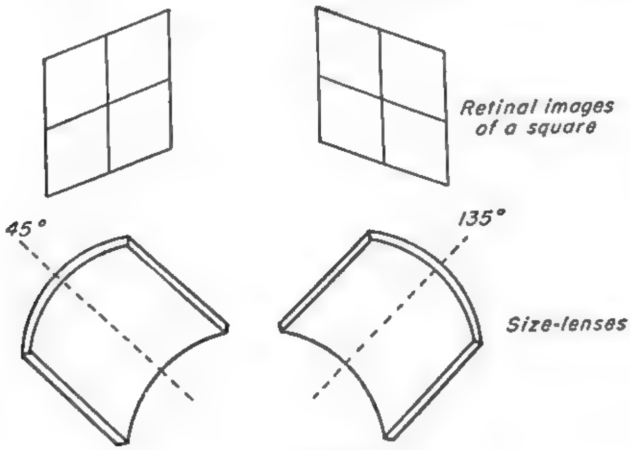


Fig. 32.

on the right, and its floor tips downward on the right. The right side wall is seen as farther away than the left. It looks to the observer as though he is no longer stationed midway between the two walls as he was in the beginning. The leaves on the

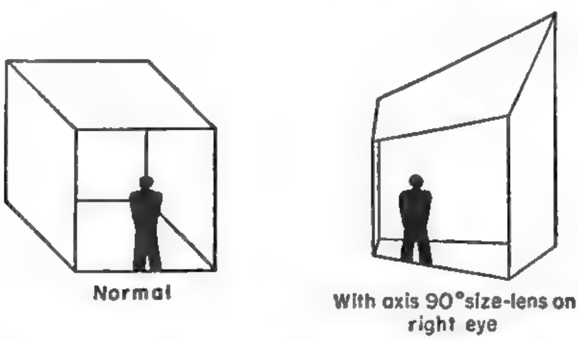


Fig. 33. Diagram to show the normally perceived shape of leaf-room (cubical), and the shape perceived while wearing an axis 90-deg. size-lens on the right eye. (See text for description.)

right side wall are larger than those on the left one, even though they seem to be farther away. The whole set of visual appearances can be reversed if the size-lens is placed in front of the left eye.

Moving Objects as Seen with Axis 90-deg. Lens. A simple way to demonstrate the binocular effect of meridional magnification in one eye is to move an object (let's say a white card) pivoted on a stick, the fulcrum of which is close to a point midway between the two eyes. If the object is moved from left to right, it appears to increase in size when the lens is worn in front of the right eye. While increasing in size, the object recedes and moves with increasing speed. The length of the stick on which the object is stationed appears to increase, and its end follows an ovoid course rather than following the arc of a circle. Retracing the movement reverses the sequence of the distortions just described.

Outdoor Situations with Axis 90-deg. Lens. There are several typical outdoor situations in which the use of an axis 90-deg. lens brings out some marked effects. If one looks down into a canyon from a bridge across it, the two walls of the canyon will tend to appear quite unequal in steepness and in distance. The canyon wall on the side on which the lens is worn will appear to be farther away and, instead of being so nearly vertical as the opposite wall, will be "flattened" or tilted much toward the horizontal. The approach toward verticality of the other wall will be much enhanced. Likewise, the wall will seem much closer. A stream at the bottom of the canyon will be changed in its rate of fall. The water may appear to be something like a waterfall or to be running uphill, depending upon the direction from which it is viewed.

In the woods, where there is an abundance of underbrush, the effects from the lens on one eye are very much accentuated. Many misjudgments occur as one attempts to make his way through a woods of this sort, and as a result the wearing of the lens on such an occasion becomes very distressing.

The Wearing of Lenses at Oblique Meridians. As was earlier pointed out, the chief effect of wearing lenses with magnification in oblique meridians is the appearance of longitudinal tilting of the horizontal surfaces, rather than their lateral tilting as in the earlier examples. Varying effects occur, depending upon whether these surfaces are grass lawns, smooth floors, or water surfaces with waves.

If lenses that make the ground or a floor appear tipped up are worn when one is in a boat and looking at the water, the surface of the water (with waves) appears much farther away than usual, and the apparent size of the waves and floating objects is greatly increased. Likewise the rate of motion of objects is increased.

With lenses that tip horizontal surfaces away from one, the waves appear to be reduced in height. With this flattening-out effect, the rate of motion of objects is decreased, and one feels quite close to the water.

Motion-sickness. Considering these effects in light of motion-sickness (seasickness, etc.), one might suppose that enhancing the subjective height of the waves would tend toward nausea when one rides in a boat. Actually the opposite effect is inclined to occur. It is the lenses that *flatten* the waves (diminish motion of seen objects) which tend to produce disturbances such as nausea and anxiety. When the lenses are removed, the untoward symptoms often clear up. One explanation of this effect is that when the waves appear flattened, the visual signals provided by looking at the waves do not "forewarn" the individual of the battering he gets while riding in the boat. Another way of stating the matter is that there is a conflict between the kinesthetic and vestibular results of being tossed and the visual perception of waves that seem too trivial to do the tossing. When one initiates his own motion by his own muscular activity, he is consistently prepared for all sorts of rapid and quickly reversing motions. When, on the other hand, one's motions are produced by external agencies, they are not always so well taken. There are not the proper compensatory muscular effects, and sickness may ensue. It is as if the tossing externally produced were some sort of a threat. Looking at waves may enable one to "understand" the tossing one is receiving. If one "assumes," however, that the waves are inducing the tossing, and if then the waves are too trivial to account for the tossing actually received, there is an inconsistency that sets up distress. In the reverse situation, one is not receiving relatively so much mechanical effect from the waves, so seeing them as large does not seem to be quite so effective, at least for some individuals.

There are several additional features of the stimulus field

that can be mentioned as factors in the perceptual outcome. Situations involving numerous rectilinear or ordinary perspective factors tend to lessen the effects produced by wearing size-lenses. The distance of objects from the eyes acts as another factor (see Fig. 34). The apparent rotation of surfaces increases as their distance from the observer is increased. There also are features of *internal consistency* that the organism cannot violate

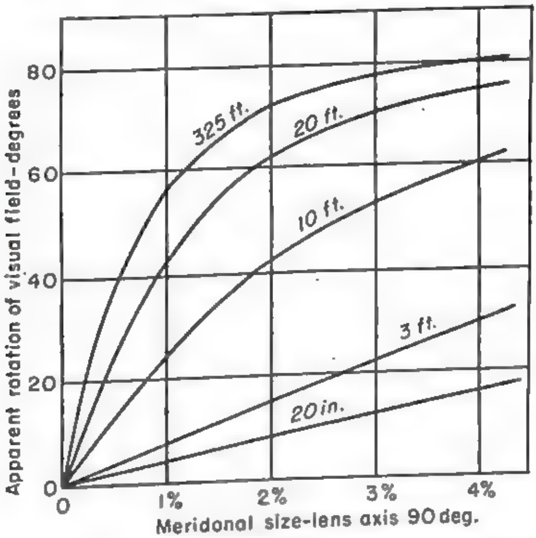


FIG. 34. Graph to show the relation between the amount of perceived rotation of an object or of visual field, and the percentage magnification of a meridional size-lens. Each curve represents the effect for a different distance. (*Ames. By permission of Amer. J. Psychol.*, 1946, **59**, 356, *Fig.* 9.)

in perception. These are complex and difficult features to analyze. For example, there are times when it seems as though a knowledge of the shapes of objects in the field tends toward their holding shape. When this is so, other features must shift in character to compensate. At other times this would-be factor of so-called knowledge does not seem to bear much weight.

Practice Effects. Practice in performing tasks and getting about in space tends to minimize the original effects that the use of size-lenses produces. No one has as yet experimentally worn size-lenses for more than a few days—at most a couple of weeks. At odd moments or in special situations during these short test-periods, the original effects became manifest, although more and

more of the time everything looked all right. Initially the constant experimental wearing of the lenses tended to produce varying degrees of discomfort, depending upon the individual wearer.

From the foregoing facts it should be obvious that the matter of space perception is a very interesting, though complex, affair. You should not, however, feel that all that needs to be known about space perception has already been found out. There are many perplexing problems, especially for the therapists who deal with eye troubles. Your generation should be able to solve some of these.

QUESTIONS

1. Why are lenses used in the study of space perception?
2. What two classes of lenses are there, and what is the optical distinction between them?
3. What may be said about the shape and other characteristics of the retinal images of the two eyes when an object is near and directly in front of an observer?
4. What changes take place in the images as the object is shifted to one side of the original line of regard, supposing the eyes, but not the head, turn to follow it?
5. Name the two general classes of size-lenses.
6. Describe the notation that indicates the axis of a size-lens.
7. In what meridian, with reference to lens axis, does the magnification of the retinal image occur?
8. Describe the perceptual effects of wearing an axis 90-deg. size-lens in front of one eye, while viewing a leaf-room.
9. What effects can be obtained by wearing a size-lens on each eye, with the axis of the two directed obliquely?
10. Give a basic rule regarding the relation of the direction of incident light and the perceived location of a point or object.

CHAPTER 13

MEASUREMENTS IN SPACE PERCEPTION

In the preceding chapter certain factors contributing to three-dimensional binocular vision were presented. A description of the effect of wearing size-lenses was the chief contribution in this presentation. Since it has been shown that alterations in the relative sizes of retinal images in various meridians distort the observer's perception of the shapes and locations of objects in space, a question which immediately arises is whether measurements of such perceptual distortions can be made.

Certain investigators set out to construct a device that would enable the making of such measurements. The device would conceivably consist in some kind of a three-dimensional target, the elements of which should be movable into various positions. When all the elements are in their "normal" positions, the target should possess a kind of bilateral symmetry. From what is already known about the distorting visual effects of meridional size-lenses used in the 90-deg., 180-deg., and the oblique positions, one should desire a target whose elements would appear in other than normal positions when the observer wears size-lenses. Ideally the target ought to be such that the wearing of a size-lens at 90 deg. would relocate the visual position of one part of it (the target); wearing a size-lens at 180 deg. shift another one of its elements; and wearing size-lenses at oblique angles manipulate still another part of it.

Since we know that the introduction of anomalous geometric relations between the retinal images of the two eyes, by use of size-lenses, results in distorted visual space perception, it might be reasoned that inherent defects or anomalies in the behavior of the visual apparatus itself might be detectable by use of some kind of a target such as we have specified. One might

also expect that since size-lenses can be used to introduce anomalies, they might be utilized to compensate for inherent anomalies when found and thus enable the individual's vision to act in line with the other psychophysiological mechanisms for enabling the individual to cope with the physical world.

On these assumptions, the target would have a use in detecting and diagnosing visual anomalies in space perception. If a person with visual anomalies of this kind becomes an observer of the ideal target we have described, he would be expected to see the elements at other than the normal positions when using the naked eye.

The first type of target element conceived of was a horizontal surface viewed obliquely and which could be tipped from side to side and tipped up at either end. This was called a *tipping board*. After some use it was found unsatisfactory for indicating specific space-distortion effects. It did not fulfill the requirements such as we have just suggested. Its faults precluded reliable readings.

You will recall that the wearing of an axis 90-deg. lens on one eye shifted the plane of a viewed surface such as that of a wall of a room (particularly a leaf-room) away from the observer on the side corresponding to the eye wearing the lens. It was thought that one might therefore utilize a plane which would pivot on a vertical axis so that its rotation could be adjusted to compensate for the effect of the lens. One might make this adjustment and determine the angular rotation in degrees, minutes, and seconds required to have the plane appear at right angles to the line of regard (*i.e.*, both sides equidistant from the observer). When this was tried, it was found that certain surface features of the plane, such as its texture, lighting, etc., tended to preclude its being a good visual target for the purpose. With a setup of this sort there seemed to be no way of separating the monocular and binocular factors used by the observer in localizing the surface. As a result, two vertical cords (plumb lines) were substituted. These appeared to be appropriate for the purpose. With these viewed against a homogeneous dark field, the observer was able to judge consistently whether the vertical cords appeared equidistant. Theoretically there are

two ways for the experimenter to make the measurement. One is for him to move the strings so that they appear equidistant to the observer and then read the angular rotation of a horizontal bar carrying them. A variation of this method would consist in having the observer himself move the cords. The second method is to provide the observer with a set of adjustable size-lenses by which he or the experimenter can make compensatory adjustments until the observer finds the cords subjectively equidistant. If these adjustable size-lens units are calibrated, the experimenter can tell what "equivalent angular shift" was necessary to make the cords visually equidistant. The usual adjustable size-lenses are, of course, calibrated in terms of the percentage amount of magnification they introduce.

Even if the technique just described might be used to measure this one form of inherent asymmetry, the asymmetry in this meridian is only one component of a possible over-all distortion condition. One would naturally reason that if the visual apparatus can manifest a functional distortion in the horizontal dimension, it might well exhibit distortions in other meridians—kinds of distortions similar to the effects produced by wearing oblique size-lenses in front of the two eyes.

In light of the several ways in which size-lenses were used to produce space distortion, other elements were added to the two-cord target. A second set of two cords of the same lateral separation were added behind the first two. Between these two pairs of cords, now forming the corners of an imaginary rectangle, a vertical cross made of cords was placed. Its elements were oblique. In some cases a vertical cord was included in the plane of the cross, intersecting its elements where they intersect each other. Figure 35 provides a clear indication of the arrangements of the elements in the complete three-dimensional target. This target, together with the reduction screen,[1] the calibrated-compensation size-lens unit, and the chin-rest has been called a *space eikonometer*. It is called an eikonometer because the Greek word *eikon* means "image." An eikonometer would be an

[1] A reduction screen is a vertical wall in which is cut a rectangular aperture just large enough to permit viewing a target behind it, but obstructing top and bottom views which would be involved in providing clues for perspective.

image measurer. It is a *space* eikonometer because these images have to do with the perception of space.

It was found that visual asymmetry, by reason of wearing a size-lens (axis 90 deg.) on one eye, or an inherent anomaly of the same type, causes not only the front pair of cords to appear unequally distant but also causes the cross to be rotated so that the side corresponding to the side on which the lens is worn will seem farther away.

It was discovered that if a visual asymmetry exists by reason of wearing a lens (axis 180 deg.) in front of one eye, or an inherent

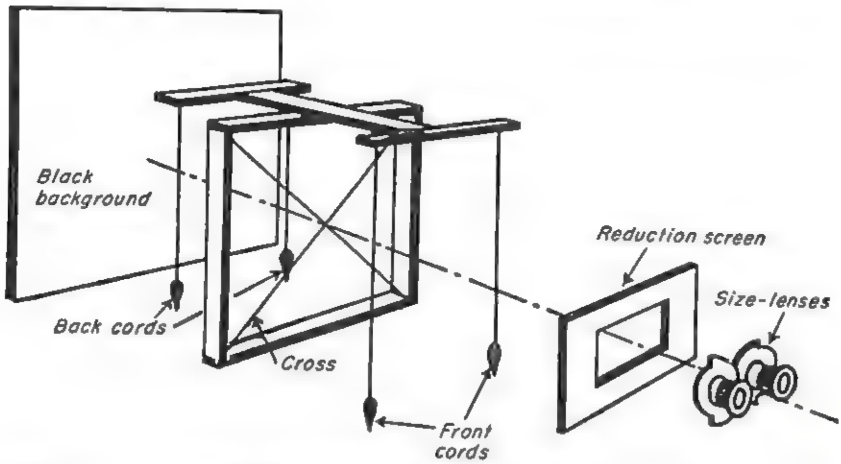


FIG. 35. A perspective drawing to show the working elements of the device called a space eikonometer. The drawing is schematic but should be quite intelligible from the text discussion. (*Ogle. J. opt. Soc. Amer.*)

anomaly of the same sort exists, the apparent position of the vertical cords will not be changed but the cross will appear to have been rotated so that the *nearer* side will now be the one corresponding to the side wearing the size-lens. By making the necessary compensatory rotation of the cross or by adjusting the 180-deg. components of the adjustable size-lens unit, the appearance of the cross can be corrected so that its elements look equidistant.

It turned out that by introducing oblique size-lens combinations, the cross could be moved into another direction. It could be tipped so that the top would appear either toward or away from the observer. The cross, of course, appears out of its nor-

mal position also when an inherent visual anomaly of the oblique sort exists.

The complete adjustable size-lens unit included two meridional size-lenses of fixed magnification geared together so that the rotation of one induced the equal but opposite rotation in the other. To begin with, in a test, the two lenses were set at the zero mark, meaning that the lens for neither eye was rotated. If the axes of the lenses are rotated in the nasal direction, the top of the cross will appear nearer than the bottom. If the axes of the lenses are rotated in the temporal direction, of course, the top of the cross will recede and the bottom will come nearer.

In order to make space distortion tests, the subject must wear his ordinary glasses so as to possess the maximum visual acuity. In any case, his visual acuity and presence or absence of astigmatism are checked. Blurred images preclude the precise interplay between images of the two eyes for the purpose of denoting spatial relations. You will recall that in the preceding chapter it was pointed out that early experiments in space perception were hampered by blurred images. This was because ordinary power lenses and prisms produced image blurring when manipulated in the ways necessary to study the effect of decreasing or increasing the sizes or distorting the shapes of retinal images. The same principle is involved in the present situation. If the subject (observer) has indistinct vision (supposing refractive error), it must be corrected before being examined by the space eikonometer.

One of the tests which antedated the space-eikonometer test was invented during World War I. The test to which we refer is the Howard-Dolman test. We shall call it a three-peg test.

The Howard-Dolman Test. The test consists of three vertical wires, pins, or pegs. Two of them are stationary, and the other one, standing between them in the visual field, is movable. The subject moves the adjustable element until it appears to be in line with the other two. The two stationary elements are a meter or two from the observer, or may actually be placed at any distance up to 5 or 6 m. They stand in a vertical plane at right angles to the line of regard and are relatively close together. They subtend, therefore, a relatively small visual angle and thus

concern a very restricted portion of the subject's space field. Whatever significance the test may have, it pertains only to this narrow range. In contrast to this, the space eikonometer covers at least 12 deg. of visual angle.

A Distinction between Monocular and Binocular Vision. Whatever usability a test has for detecting the nature of spatial localization is based upon its ability to distinguish between monocular and binocular clues. Some third-dimensional seeing is done by use of monocular vision, that is, without the aid of the clues that differences in the two retinal images provide. Monocular vision involves the use of visual angle, movement, parallax, overlay, relative indistinctness of contour, and a number of other factors including the differences between objects of known and unknown or ambiguous nature.

It is for this reason that a brief consideration of the distinction between monocular and binocular vision is given by way of an illustration. It will enable you to appreciate one of the reasons why strings, wires, or cords are better as target elements than objects with appreciable surface per unit length.

Targets can be viewed by use of one or both eyes. The use of only one eye precludes the organism's employment of certain mechanisms that may be brought into play when two eyes are used. Although one-eyed vision is by necessity monocular vision, two-eyed vision is not necessarily in the full sense binocular vision. Binocular vision, as might be supposed from what has already been said, must involve some utilization of the differences in the shapes, sizes, and locations of the retinal images of the two eyes for perspective-forming purposes.

To illustrate a distinction between monocular vision and binocular vision, let us examine a two-element target, such as is illustrated in Fig. 36. It consists of two vertical pegs that stand behind a reduction screen. If one views these two pegs with one eye, one gains the impression either that the pegs are equidistant from him or that one peg is farther away than the other one. The ways to make one look farther away than the other are (1) to place one peg farther away, as in B of Fig. 37, or (2) to make a difference in the diameters of the pegs, as in C. What is actually done in both cases B and C is to make the visual angle subtended

by one of the pegs less than that of the other. With monocular vision this difference is generally effective for placing them at unequal distances visually.

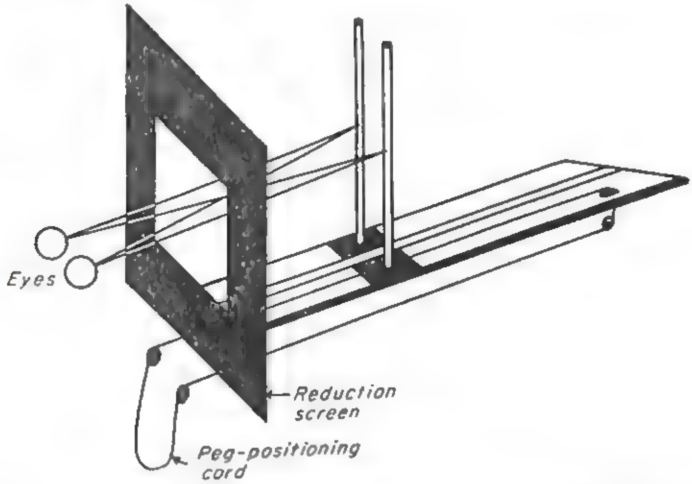


FIG. 36. The arrangement of a two-peg device for judging distance and indicating the difference between monocular and binocular vision. (For further explanation consult Fig. 37.)

Now let us examine what would happen were the subject allowed to use two eyes, and presuming he has good binocular vision. When he looks at A with both eyes, the pegs should appear equidistant. When he looks at B, the one peg produces

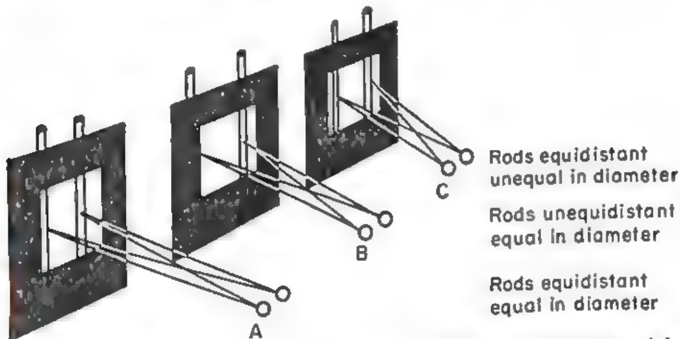


FIG. 37. Three diagrams to show three possibilities in the positions and sizes of rods lying behind a reduction screen. Monocularly, in A the rods should look equidistant; in B and C, unequidistant. Binocularly, the rods in A and in C should look equidistant, but in B they should look unequidistant. Thus spatial location is better perceived by binocular vision.

a large image on both retinas, and the other produces a smaller one. There are differences in the sizes and in the directions of the images of the two pegs, which was not true when the pegs were actually equidistant. The images do not fall on symmetrical parts of the two retinas when one looks midway between the two pegs. When one looks at C, the same symmetry exits as was true for A, but now only the sizes of the retinal images of the two pegs differ. Whereas this total retinal situation is taken to signify a difference in diameter of the two pegs, the retinal situation when looking at B was taken to mean a difference in the distance of two equisized pegs.

This is the reason why fine wires, rather than pegs that present a surface to the eye, are better to use as elements in a three-peg test or any other test for binocular space perception. By using fine wires or cords, the potential *conflict* between monocular and binocular space perceptions are obviated. To show that there is a conflict, or at least the occasion for a possible compromise when using surface-presenting elements such as pegs, the following will serve.

One can calculate the different visual effects that would accrue from one-eyed viewing, and from two-eyed viewing when truly binocular, on the basis of what has been described above. Then one can try the two kinds of targets (B and C) upon subjects. In doing so, one will find that some of them get the wholly binocular effect; some approximately the monocular, when using two eyes. Others report an effect somewhere in between the two effects. That is (in B), they localize the farther peg somewhere between where it should stand and where it would be if equidistant with its fellow peg. In such cases this is taken to indicate that neither monocular nor binocular clues were fully dominating. Instead, some sort of a compromise was in operation.

A very great number of problems in space perception cluster around those already presented here. What has been given, however, should form a beginning demonstration of the type of experimentation possible and necessary in space perception. The chapter to follow will take up another but totally different aspect of connection between the individual and the physical world in which he lives and must operate.

QUESTIONS

1. Describe the space eikonometer.
2. What was it constructed to achieve?
3. What are two of its antecedents?
4. Why were thin cords used for the target elements of the space eikonometer?
5. What is the Howard-Dolman test, and what are some of its limitations?
6. Describe a setup whereby a distinction between monocular and binocular vision can be demonstrated.
7. What operation can be substituted for the use of an adjustable size-lens in a space eikonometer setup?
8. What is the relation of visual acuity to observing the space distortions brought about by size-lenses?
9. What effect does habit have upon perceiving spatial distortions with size-lenses?
10. Do all the symptoms of difficulty for the organism have to be exclusively visual when wearing size-lenses or when suffering from an inherent size imbalance in the two eyes?

CHAPTER 14

THE AMBIGUITY AND EQUIVALENCE OF VISUAL STIMULI

The aim of this chapter is to demonstrate certain relations between objects, their retinal images, and the resulting perceptual experiences. These relations may change from instant to instant, as a result of the object's movement to or away from the eye, changes in the object's orientation with respect to the eye, or the contribution the organism makes in "interpreting" the stimulation that it receives.

For example, a given object may be seen as one kind of a geometric pattern—a plane surface at right angles to the line of regard, or another shaped figure at some other angle of orientation to the eye. This lawful possibility of an object being variously perceived—seen as one or another of several objects—is based upon the fact that there are degrees of freedom in geometry. Viewed from the observer's standpoint, some stimulation is ambiguous. This is not all. A host of differently shaped objects may be perceived as the same one. This, too, is an outcome of the geometry involved in the situation. That geometry provides for this may be spoken of as its *equivalence*.

Although perception may be expected to express considerable freedom, it is not expected that any freedom stems from the very nature of geometrical optics itself. Geometry is generally looked upon as representing rigidity of relationships. There is a common tendency to assume that when a specific perception is at variance with what is otherwise known about an object, this discrepancy arises solely out of the nature of perception. This is not the case.

Our interest is to describe the factors in the geometry of the physical situation that produce this freedom. The simplest

example we can take to introduce the subject of equivocality, or ambiguity, or the freedom of the organism to select between alternatives, is given in the following.

Visual Angle and Object Identification. Let us say we have a darkroom in which the only thing that we can see is an object onto which some light is projected. The object's location, size, and shape are determined solely by the light reflected by the object itself. There are no other objects by which to judge its location or with which to compare its size. There is no floor or ground that extends toward the object. The object stands out alone in the darkness.

Let the object be a square white card 5 by 5 in. placed at right angles to the line of regard. To you it will look like a square. If we turn it so it is oblique to your line of regard, you will not know that it is a square, for its edges will form a trapezoid or trapezium from where you are. The sole geometric determiner of the object's seen shape will be the shape of the surface it presents to you. It will be bounded by a different set of visual angles than when in the original position at right angles to you. If we turn the card nearly all the way around so that only its edge is presented to view, it will be a line to you instead of a square or any other extended surface. Although we can say that the external object is the same when judged by other criteria, we have to say that it does not always form the same shape or size of image on the retina and cannot always be identical as a source of stimulation.

Now, let us turn the square back to its "normal" position at right angles to your line of regard. Let's imagine fine, invisible threads running from the corners of the square to your eye. These threads represent the lines which mark off the visual angles subtended by the square. If an experimenter substitutes a small square for the original one and places it nearer to your eye, it will be seen as the original square and at the original distance from the eye if it subtends the same visual angles as the first one. Note that we have not violated any geometrical principle in saying that the two different-sized cards are taken by the observer to be the same card. We have simply utilized the principle of maintaining fixed visual angles. When visual

angles remain constant, the retinal image remains fixed and the organism has no basis for experiencing a shift from one object to another.

It is obvious from Fig. 38 that it is possible to substitute any one of an infinity of plane surfaces that will subtend the same visual angles at the eye, as did the original square. Not only is this true, but an endless series of variously curved surfaces could be substituted one by one, and yet the object would continue to be seen as the same one.

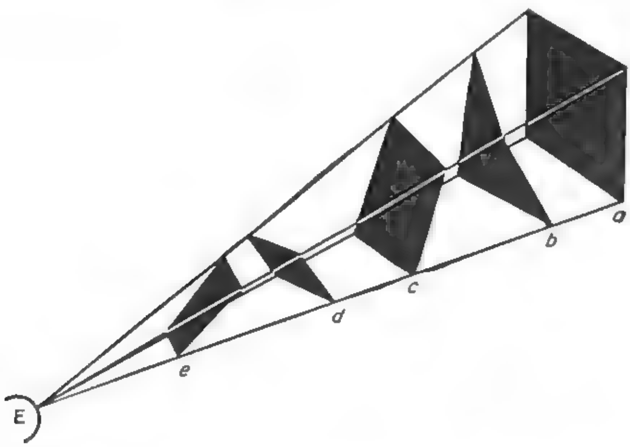


FIG. 38. A diagram to show that an unlimited number of plane objects may subtend the same visual angles at the eye. These objects, accordingly, differ in distance from the eye, in actual size, in shape, and in orientation with reference to the eye or line of regard. (See text for further discussion.)

A vivid example of a case in which various-sized objects are seen as the same size occurs in the following situation. Let a playing card be the only thing seen in a darkroom. This condition, as in the case of the white card previously mentioned, is brought about by spotlight projection of light onto it.

An observer can be asked to note or tell the distance of the playing card. Then, when the spotlight is turned off, another playing card is substituted. This time the card is a giant replica of the first one, and is placed at the same distance as the first one. Again the observer is asked to tell how far away the card is. His report indicates that he sees it as being much nearer to him than the card in the previous exposure. In fact, to him

it is merely the same card brought nearer. In the third trial a miniature replica of the playing card is shown to the observer. This time he indicates its position as much more distant than in the first trial. Here we have, then, three objects literally taken to be the same one, but in different positions.

Surface Structure and Object Orientation. If, however, various surfaces subtending equal visual angles contain certain types of internal differentiations such as stripes or cross hatching, the

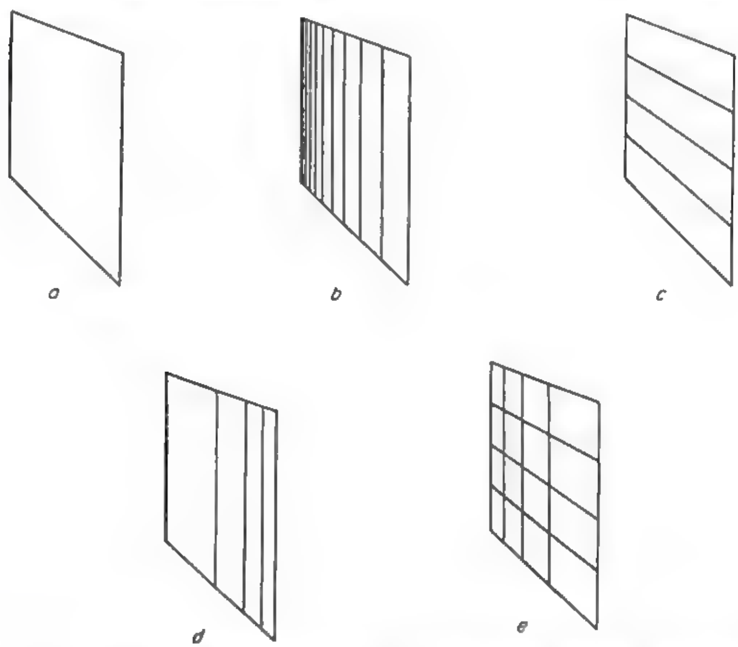


FIG. 39. A series of geometrical figures all identical in over-all contour but variously differentiated (structured) within. Accordingly, their orientation from the observer is different and their perceived shapes are also different.

freedom of visual perception already described is somewhat reduced. The principle is illustrated in Fig. 39. Surface *a* may be oriented either at a slant or at right angles to the line of regard, as far as visual perception is concerned. Again, differences in visual orientation are possible, as illustrated in *b* and *c* of the figure. In *e*, however, there are additional factors tending to make the surface appear as a rectilinear one whose one side is considerably farther from the eye than the other. If, as in *d*, the spacing of the vertical stripes or lines is not such as would

occur in a natural situation when a surface with equispaced elements was seen at an oblique angle, then the rotation of the surface from normal is not so likely to be perceived.

Equivalence in Three-dimensional Objects. One of the best demonstrations of the equivalence of various objects is provided by an arrangement constructed at the Dartmouth Eye Institute. At eye level, for the standing individual, is a peephole in a vertical screen. The observer is asked to look through the hole and tell what he sees. He answers that he sees a chair. This "chair" that the observer sees is produced by looking at a three-dimensional object consisting of white wires arranged in three-dimensional space in the form of the usual straight-backed chair. The wires simply form the outline or "frame" of a chair, as is shown in Fig. 40.

Naturally, this object is a chair as viewed from any point and as would be felt by the hands were the opportunity afforded. There is, therefore, nothing unexpected in the observer's saying that he sees a chair.

In the vertical screen bearing the first peephole (P_1), there is a second one (P_2) about 2 ft. to one side. The observer is asked to look through it and tell what he sees. Again he says he sees a chair. This chair is in every respect the same as the first chair and may even be the same one for all he knows. It is the same size and similarly oriented with respect to his regard and it is the same distance from the eye as the chair seen through the first peephole.

The "object" which has provided visual stimulation in this second trial is very unlike the first chair. If we examine it from the rear of the setup, where we can detect its construction, it will be found to be a two-dimensional structure. It is a distorted horizontal two-dimensional frame of white wires so arranged that from the peephole its elements subtend the very same visual angles as the first "chair." This is also illustrated in Fig. 40, in which a "chair" is drawn on a horizontal plane so as to subtend the same visual angles as a three-dimensional chair.

The observer is asked to take a second look at the first chair from the first peephole and then view the second chair again so

as to be sure to see if they are the same. As a result, the observer assures the demonstrator that they are identical.

A third peephole (P_3) is to be found still farther to one side of the initial one. The observer is asked to look through it. And,

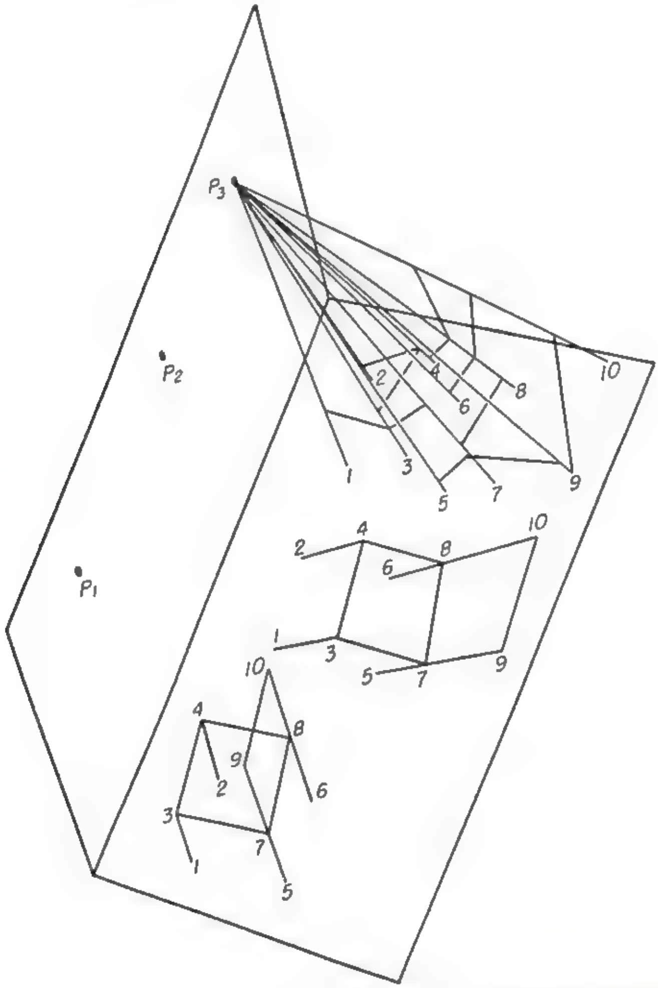


Fig. 40. A diagram to show the elements of an arrangement used to show that three "objectively" different physical constellations ("objects") may give rise to identical retinal patterns and therefore be seen as the same object. (*Ames. Dartmouth Eye Institute.*)

as in the previous trials, he is to tell what he sees. Again he reports he sees a chair that is exactly like the one seen through the first two peepholes.

The "chair" in this case is an arrangement ("object") very much *unlike* the first two. It consists of a number of white wires irregularly strung from a group of black (invisible) fine wires whose near ends are attached to a ring immediately surrounding the peephole. The far ends are anchored in a horizontal black plate. These wires are so arranged as to form the visual angles of the various corners of the chair seat, chair back, and chair legs. The white wires, just so long as they are strung from the proper black wires, form the contours of the chair. This is true, regardless of what plane they lie in with regard to the line of sight. Thus some wires can be short and others long, obliquely oriented.

When the observer has viewed each of these "chairs" alternately several times and has assured himself that they are perfectly identical, he is shown their construction. Only when they are viewed through the proper peepholes are the last two objects chairs. But in that case they are unquestionably and unambiguously seen as chairs.

The point to be brought out by this demonstration is that visually one "object" is just as truly a chair as either of the others. The greatest significance of a setup of the kind just described is to be grasped only when one serves as an observer and experiences the results himself. It is then that the situation induces a *value* experience—in other words, evokes an emotional response. It is in this direct experience that these setups begin to take on a specific set of personal meanings. They should not be regarded as trick arrangements, but as demonstrations of principles which everyone should be vividly aware of.

Space Distortion and Emotional Consequences. The demonstration very vividly suggests the possibility of contradiction between senses and thus the production of conflict. You may feel, however, from reading about it, that the demonstration has shown that anomalous visual perceptions can be and usually are quickly corrected by use of the sense of touch or some of the other senses. This is not necessarily the case. Visual perception may, under some conditions, dominate all others. The fact that one sense does not always completely outweigh the others is significant from the standpoint of conflict, hesitation, and the loss of self-assurance.

An example of the emergence of uncertainty and uneasiness, not to say fear, is to be had in the use of size-lenses described in the previous chapter. The first time one puts on a pair of spectacles containing an axis 90-deg. size-lens and goes outside and attempts to walk across a street, he will immediately feel very unsure of himself. He feels incompetent and unsafe. He is in a new world. Such is the experience, even though he asks someone to guide him across the street. The visual impressions that he receives, even though known to be anomalous, are the only visual basis for experience and action that he has, and he feels compelled to take them at face value, even though they are not very useful in accomplishing his immediate purpose. The convincing and compelling features of sense perception are two of its salient attributes.

The foregoing description has been intended to suggest the possibility of the fruitfulness for psychological study in this area. In comparison to what we need to know, experimental work in this area is relatively scant. An approach starting from an appreciation of using the organism, as well as its surrounds, as a reference can be expected to be quite fruitful in obtaining information of a very helpful nature.

QUESTIONS

1. What is the aim of this present chapter?
2. What is meant by equivalence?
3. What is a "visual angle"?
4. What may be said regarding the appearances of all two-dimensional objects that subtend equal visual angles when such objects are the only objects seen in the visual field?
5. What does geometrical equivalence lead to in perception?
6. What does the chair experiment described in this chapter indicate to you?
7. Is the term "illusion" a scientific term? Justify your position.
8. Are the terms "stimulus" and "object" synonyms? Justify your answer.
9. Describe a situation in which the wearing of size-lenses is likely to induce uneasiness or a lack of security.
10. If all three setups in the chair experiment are chairs — why? If not, why not?

CHAPTER 15

VISION UNDER OPTICAL MAGNIFICATION

We have not finished our survey of the spatial attributes of seeing and the conditions on which they rest. One of the important areas of knowledge about vision involves what happens to the appearances of objects when the observer uses optical magnifying instruments such as binoculars. Various amounts of optical magnification are commonly used. These range from the 2- or 3-power opera glasses, 5- to 8-power bird and field glasses, to the 7- to 10-power hand-held military binoculars.

Certain distortions in the visual field are bound to be produced by all binoculars. The amount and obtrusiveness of the distortions depend upon the magnification afforded by the instruments. The spatial distortions to which we refer are not dependent upon optical defects either in the lenses of the instrument or in the observer himself. The distortions to be described are the results of certain inherent limiting features in the way magnification is brought about by binoculars. To anticipate our story, it may be said that whereas, in unaided vision, the size of the image on the retina is increased by increasing the size or nearness of an actual object, this is not the case in vision with binoculars. With binoculars, the size or the nearness of objects is produced solely by enlarging the images on the retinas, while the object and observer remain at fixed distances from each other.

Moving toward or away from an object causes a shift in the size of the images upon the retinas, in which case all parts of the images are altered in due proportion. As a result no perceptual distortion results. Changing the sizes of the images upon the retinas by means of binoculars does not shift the sizes of all components of the images in the proper proportion. Inspection of Fig. 41 will indicate that certain dimensions of the image

are magnified (those at right angles to the line of regard), while others (those in the direction of the line of regard) are not magnified. This produces a flattening of the object-as-perceived. This disproportionate manipulation of components of the retinal images with the concomitant visual distortions is the main concern of this chapter.

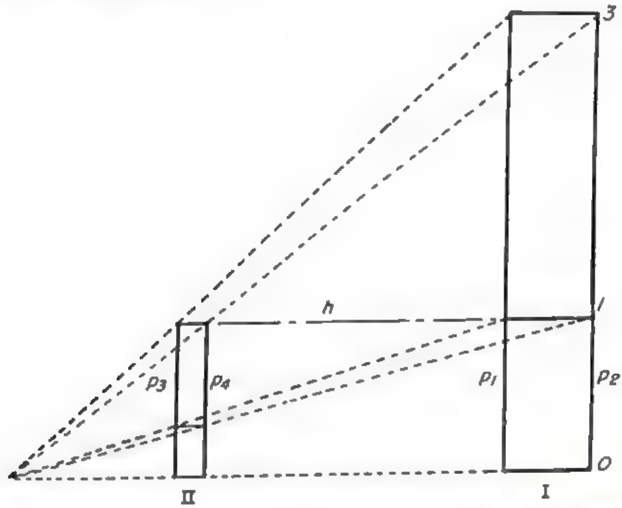


FIG. 41. A diagram to indicate the principles involved in magnification with optical instruments. The actual size, the magnified size, the actual distance, and the decreased distance of the object under magnification of the retinal image are shown. (For a further understanding of the principles consult the text.)

Our preliminary considerations will best pertain to reviewing what binoculars do that is so desirable in many situations. First of all, they enlarge the retinal images of the objects viewed. In doing this, they are not able to include as much of the field in front of the observer as he can take in with the naked eye. The two essential types of binoculars differ considerably in this respect. The older or Galilean type provides a very restricted visual field, and the newer or prismatic binoculars include a visual field two or more times as great in lateral range.

There are two major factors involved in the use of a magnifying instrument such as a pair of binoculars. The one is the *light-gathering function*, and the other is the *magnification function*. The light-gathering function has to do with improved visibility

in low illuminations and, of course, is not actually independent of the enhancement in size of retinal image brought about by the magnification function. Not all glasses (binoculars) of the same magnifying powers gather equal amounts of light. Those that gather the most per magnification are called night binoculars. These have larger exit pupils,[1] and to have them, the glasses must be larger in other dimensions.

There is some loss of light within the glasses themselves, but in spite of this they gather a greater sum total of light into the image of an object than is provided by unaided vision. There are decided practical limits to increasing the diameter of the exit pupils, but even aside from this it would be useless to increase the exit pupils much beyond the size of the natural pupils of the observer's eyes. The human pupil opens to a maximum of approximately 8 millimeters in diameter in total darkness.

The effect of gathering more light into a retinal image of an object is to increase the visibility of the object. The object will thus either be visible when otherwise it would not, or else it will appear brighter than it otherwise would. This improvement is brought about despite the fact that for each unit area of the retinal image there is a little less light than with the naked eye. The human visual mechanism is a summating device and therefore can take advantage of the *increased area* of the retinal image and translate this into increased brightness. This utilization of area is nothing peculiar, of course, to the use of magnifying instruments. In unaided vision area affects the level of the threshold and is a factor in determining the sensation of brightness. Objects subtending large visual angles have a lower threshold than those subtending small ones. Thus in their light gathering function we have run across nothing unique about the way binoculars behave, or as we might say, how we perceive when aided by binoculars.

This is not the case in regard to the magnification function. Several unexpected results show up. There are two matters that require our attention. They are (1) the consideration of the possible differences in magnifying flat surfaces (two-dimensional

[1] Exit pupils are the apertures through which the light leaves the binoculars to enter the eyes of the observer.

objects) and three-dimensional objects; and (2) the question of *where* objects are seen when magnified.

The second of these will be dealt with first. If you have observed carefully and can remember the effects when using an opera glass or a bird glass, you will recall that the major visual effect was to make objects appear closer than with the unaided eye. Birds did not seem five times their natural size just because your glasses magnified the retinal image to five times its original diameter. If you looked at a visual field containing nothing but objects of unknown size, you might see them as larger than you would were you to look at them with the naked eye. In fact, certain scenes which you will view with your bird glasses may contain objects that are brought somewhat nearer than with the naked eye, but are also magnified somewhat. Such a result is a compromise between *procession*[1] and *magnification.*

It should be pointed out here that the procession and the retention of original size by objects is an example of a common effect generally called *size-constancy* (see Chap. 17). Naturally, a size-constancy would be expected to apply more nearly universally to objects that in everyday life maintain a fixed size or narrow range of variation. Even in the case of a range in size, the organism reacts as though it were an "averager." For example, an observer is able to call a specimen oak leaf either "big," "normal," or "little." Better still—in certain distance-judging experiments, the location of a leaf specimen will be unwittingly judged as if it were an average-sized leaf.

Spatial Effects of Magnification. Assuming the fact that an object is visually brought closer to the observer, we can now examine what takes place in magnification. To do this, it will be well to inspect Fig. 41. The diagram in this figure illustrates the actual position and size of the object, its enlarged size, and its position seen with the aid of a 3-power field glass. The actual position is indicated at I, and the newly seen position at II, one-third the distance from the eye.

Lines are drawn to indicate the visual angles subtended by both the back corner and the front corner of the object. These

[1] Procession is to the word proceed as recession is to the word recede Proceed, in the present use, is to come closer.

are projected into the eye to indicate the size of the retinal image. For the actual-sized object the angles subtended are suggested by $0 \longleftrightarrow 1$. For the "magnified object," they are $0 \longleftrightarrow 3$. Broken lines are drawn to picture the visual angles involved.

Since the object is actually seen at about one-third the distance of the object seen by the unaided eye, the "seen" object is placed in accordance with this in the diagram. Naturally, it is construed as lying within the visual angles belonging to the "magnified object." Thus to complete the diagrammatic construction of the object as seen, it was proper to use the visual angles of the magnified retinal image and erect perpendiculars to the horizontal h indicating the natural height of the unmagnified object. When these perpendiculars, p_3 and p_4, are erected, they form the near and far edges of the seen object. It will be noted that they lie closer together than p_1 and p_2. This means that the object possesses a decreased third dimension (decreased perspective). Three-dimensional objects then ought to look flattened. This is actually what happens when one looks through a pair of binoculars. Three-dimensional objects do become flattened. In fact whole visual fields become flattened.

A way to show this is to use a cardboard rectangle 10 by 20 in. that is folded into two equal halves, 10 by 10 in. This can be placed on a table in a position representing an open book with its back to the observer, as illustrated in Fig. 42. Using a 3-power glass, let an observer look at the cardboard. If the cardboard is bent at right angles, it will look as though the angle of the fold is much less. That is, it will look as though the card were opened and more nearly flat, as in the broken lines in the figure. A quantitative experiment can easily be made out of this simple setup by arranging it so that the observer can adjust the angle formed by the two halves of the card so that it looks like a right angle. Obviously, the angle will have to be made much more acute than a right angle to look like a right angle. The observer might vary his distance from the card to see the possible effect upon the apparent angle of fold.

Unexpected Rate of Visual Movement. Another aspect of visual perception is greatly distorted by the use of binoculars, and that is visual motion. If you look down a highway while

wearing binoculars and watch a car coming toward you, you will note that the car moves most amazingly slowly. On second thought, this is actually what is to be expected, if one considers the effects we have already been discussing. First of all, the car looks near to begin with, even though it is actually far away. Since it is far away, it must traverse a great distance to reach you. Because you see it as being near to start with, it would psychologically follow that to reach and pass you it would take

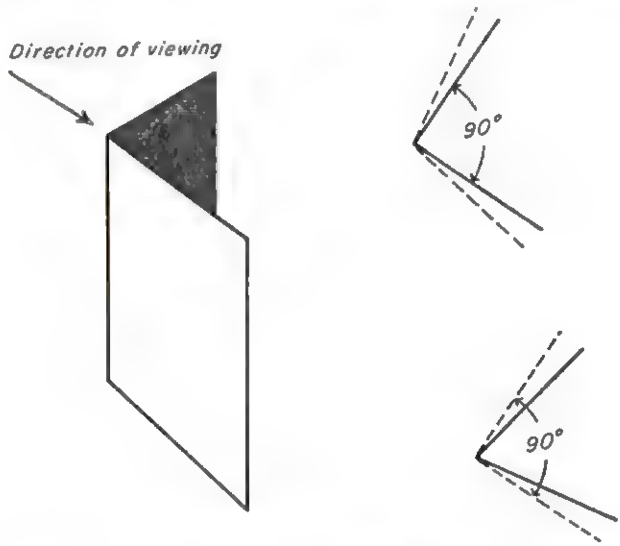


FIG. 42. Diagram to indicate the flattening effect on three-dimensional objects when viewed through optical magnifying instruments. A card is bent in the middle. If the two parts form a right angle, they are seen to form a more obtuse angle. To be seen as forming a right angle, the card must be bent to form a more acute angle. By manipulation of the amount of bending needed to produce the appearance of a right angle, one can measure the flattening effect.

very little time. But since it actually requires considerable time to do that, you see it as moving like a slow tractor toward you. Even though you understand this, the car still requires a longer time to reach you than you expect. Tacit in this expectation is, of course, the premise that cars travel rapidly—at least much more rapidly than the rate at which one walks or the slow pace seen through the binoculars. Your expectation is in accord with your knowledge of usual car speeds.

This distortion of the temporal aspect of motion is but another example of the distortion of the third-dimensional features of seeing. Timing of lateral motion is not quite consistent with what one sees, but there are several limiting conditions precluding the inconsistency's being noticeable.

Chinese Perspective. Another example of the change in apparent shape of an object is brought out when one views a more complex object, such as a church steeple or clock tower, through a pair of binoculars. If the observer looks carefully, he can verify the effect shown in Fig. 43. In the left cube shown, the opposite sides are all parallel to one another. In the next cube, which is complicated by the addition of a set of lines that would make it look like a box with a lid, magnification produces an effect known as Chinese perspective. In this form of perspective the upper and lower edges of the object do not appear to be parallel. In keeping with this, the edge near the observer is shorter than the far edge. This is the effect when one views a clock tower, for example. Its near edge is foreshortened, and the expected horizontal parallels markedly converge toward the observer. With the naked eye this does not happen. In all cases, the clock tower, or the box with a lid on it, resists being flattened out of its rectilinear form. It must therefore look the way an object would if it were not "flattened." In behaving this way, it must function as an object producing a retinal image such as the tower or box actually do. By considering the geometrical optics of the situation, it will be found that the object would be one with the foreshortened shape already described—one with shortened near corner and nonparallel upper and lower edges. This can be shown by graphic projection methods using principles such as were used to account for the flattening effect in the cube object, or the folded cardboard (see Fig. 41). Chinese perspective is the outcome in a situation seeming to offer alternatives to organism. It is simply illustrative of one of the endless number of situations in which the organism is faced with alternatives. In some types of situations the alternatives are about equal, and perception fluctuates from one alternative to the other, as in viewing *reversible-perspective* figures. In other situations the organism holds tenaciously to one or the other of

the alternatives. Just why it accepts the one alternative rather than the other is not always apparent.

The general rule is that whatever the perception resulting from the retinal image, it will possess a kind of self-consistency. This consistency stems from a specific set of factors that, if the process

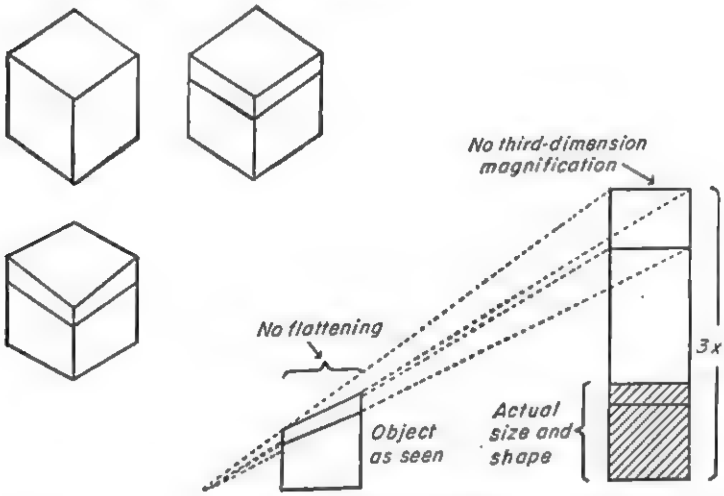


Fig. 43. Diagram to show what is called Chinese perspective, *i.e.*, the foreshortening of nearby elements in a representation of a three-dimensional object. Complex objects undergo this change in appearance when seen under magnification by optical instruments. The same scheme in picturing the effects is used here as in Fig. 42.

were *thought* instead of perception, we would call *premises*. The retinal image is not "ignored," it is utilized; but since it is not something that requires a single end-result, it is the foundation for even the peculiar perceptual end-results which we sometimes encounter. To follow out and account for the results we attain by experimenting with perception we have to be aware of this rule and be insightful in discerning the applications. When we do, we find that it provides us, step by step, with demonstrations of the intricate self-consistency of the perceptual processes.

QUESTIONS

1. What is the use of studying vision under optical magnification?
2. What are the factors that vary the size of the retinal image in naked-eye vision?

3. Does optical magnification always make objects look larger?

4. When does optical magnification make objects look nearer?

5. What two principal functions are involved in the use of optical magnifying instruments?

6. What is Chinese perspective?

7. What effect does optical magnification have upon the apparent movement of objects?

8. What is meant by the flattening effect when involved in the use of optical magnification?

9. Describe a simple way of measuring the flattening effect.

10. What is meant by the statement that "premises" are involved in perception as in thought?

CHAPTER 16

LIGHTNESS- AND COLOR-CONSTANCIES

The human organism treats stimulation in a fashion meaningful to it. Although we are in a universe that at bottom is thingless, consciousness is comprised of our own constructions—*things*. The aim of experimental psychology is not to study the physical world, but to understand the human individual's phenomenal world of objects in relation to it.

The perception of *things*, or objects, involves certain operations on the part of the perceiver. Among them are the maintenance of thing-identity, thing-continuity, and thing-stability. The tendency to maintain this object stability or uniformity is commonly spoken of as *constancy*. The matters that we wish to discuss in this chapter pertain to two aspects of constancy.

This area covers not only the principle of constancy in the broadest sense but also principles of contrast and similarity. Since vision is the best modality for our purpose, we shall use it for our experimental illustrations.

We shall consider first the phenomenon of lightness-constancy. The simplest example is to be had by using a sheet of paper that looks white in daylight. If you place it in a deep shadow somewhere, it continues to look white. It will be found that at night under a reading lamp it still appears to be white. This is to say, it retains its identity largely despite the amount of light reflected from it to the eye. It can be distinguished from a sheet of gray paper under each of the conditions just mentioned. The means of retaining its identity is, as just seen, the retention of the apparent property of whiteness. Take a sheet of white paper and a sheet of gray paper, as judged by how they look in daylight, and put the white paper in a deep shadow and the gray paper in intense illumination. When you look at them for com-

parison, you may find that the white paper looks white and the gray paper looks gray. This will happen in spite of the fact that the white paper reflects less light per unit area to your eye than the gray paper does.

If we use the arrangement shown in Fig. 44 (left), the matter can be shown still better, largely because we can manipulate the variables in a known quantitative way. The arrangement con-

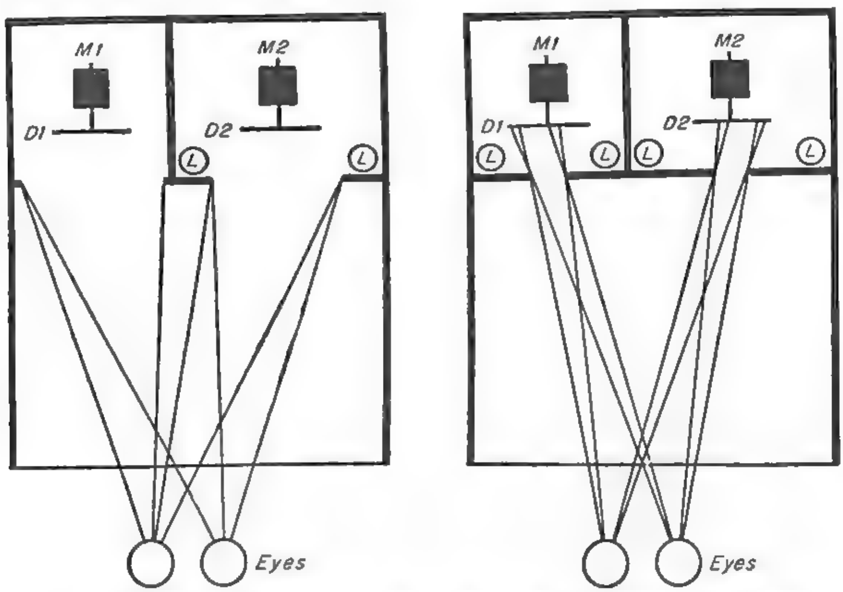


FIG. 44. Diagrams to indicate certain principles of perceived lightness of object surface. The surfaces in the left-hand diagram are those of disks D_1 and D_2, as seen against the backgrounds of the compartments in which they are located. In the right-hand diagram the disks reflect light through square apertures of reduction screens and therefore are not seen as disks, but as squares of gray. (See text for further discussion.)

sists in a box with a dividing partition so that one of the compartments can be illuminated and the other left in partial darkness (in "shadow"). M_1 and M_2 are motors operating ordinary color-mixing wheels. D_1 and D_2 are disks carrying black-and-white sectors of paper adjustable to form any desired surface ratio of the two. With these disks (color wheels) various grays can be produced when the wheels revolve at high speed. One can set the disk D_1 at some fixed black-white ratio and thus have a fixed standard of gray for comparison with a gray on disk D_2.

When both disks are in equally dark shadows, one would expect that the black-white ratios on both disks would have to be approximately the same for the two disks to match in grayness.

After observing the two revolving disks under these conditions, we can turn on the light in the right-hand compartment. If there were no such thing as lightness-constancy, the right-hand disk would look many times lighter than the left-hand one. Actually the two disks may look equally white. If they do not, they may be made to match by increasing the amount of black relative to the amount of white on the right-hand disk. The differences of the ratios of black to white on the two disks would then be a measure of the effect of using the higher illumination in the right-hand compartment.

If complete constancy were to exist, the original black-white ratio on the right-hand disk would not have to be changed. If constancy were only partial or approximate, then a small change would be required. The amount of this ratio change is the measure of the approach to constancy.

Were we now to place a reduction screen in front of the revolving disks, a new set of viewing conditions would be established. Figure 44 (right) indicates the new arrangements. What the individual now sees is two squares of gray. There is no longer anything to indicate to the observer that he is looking at revolving disks, or surfaces separate from the walls of the compartments. Under these conditions varying the illumination of the right-hand disk will vary the lightness of the right-hand gray square accordingly. There will be no tendency whatsoever for the maintenance of lightness-constancy.

These demonstrations bring us to the supposition that there is something about the visual field of which the comparison areas (sheets of paper, or disks in our case) are a part that helps determine the outcome. Whatever these factors are, they contribute to the determination of whether the light involved in the visual target in question is considered as *emitted light* or as *reflected light*. In the case of emitted light, the nature of the emitter is not in question. The observer is not attributing the light involved to the properties of the emitter as a *thing*. The identification of a thing by the light coming from it does not operate in such a case.

In contrast to this mode of reacting to light, there is the response to light as a reflection from an object. In this case the identity of the object is involved. Since the surface in question is part of a thing, it tends to retain its original properties. It is also a rule that light reaching the eye as by reflectance is perceived as more intense than when perceived as mere luminance. When light is perceived as reflectance, it is seen as lightness; and when perceived as mere luminance, it is seen as *brightness*. Constancy has to do with *lightness* rather than brightness.

Our topic includes also the question of perceiving *color* as an aspect of light. We speak of achromatic colors, or the grays, and of chromatic colors, or the spectral colors, such as red, orange, yellow, green, blue, violet, and their intermediates. Colors may be seen either as *surface colors* or as colors whose source is otherwise or less well defined. Surface colors are sometimes called *object colors*. From what we have already said, it would be expected that color seen as a property of an object tends to maintain some degree of constancy. Colors whose source is less determinate would be expected to behave somewhat differently. Color includes white, broadly speaking, as well as the spectral hues. Since hue and lightness of objects are quite intimately bound up with each other, they will not always be strictly separated in what is said about constancy.

One of the situations that illustrates some of the principles already alluded to is the following. Let a sheet of black (*i.e.* poorly reflecting) cardboard be hung in the open doorway to a dimly lighted room. Let the light from a projector be cast on this cardboard. When the light falling on it is intense, the cardboard will tend to appear white. If a small piece of white paper (a good reflecting surface) is interposed between the cardboard and the projector, the cardboard's remaining visible surface now looks black. The cardboard looked white originally because it was the only visible surface receiving the projected light. No comparison surface was available, therefore the total amount of the light reaching the eye was the crucial factor. When the white paper was introduced, and it too received the intense illumination, the real color of the two surfaces was then perceived.

When the black cardboard is seen alone, reducing the inten-

sity of the projected light upon it by definite steps results in the perception of a number of shades of gray throughout a range from white to black. If the cardboard is accompanied in the visual field by a highly reflecting surface (white), then the manipulations in the intensity of the projected light will not alter the colors of the two surfaces. The cardboard will remain black, and the other surface will remain white throughout the range. The two surfaces exemplify constancy. The factor that changes, for the observer, is the *illumination*, not the colors of the surfaces.

One may manipulate two visible areas somewhat better with a two-projector setup. The slide used in the first one is opaque except for a ½-in. circular hole. Thus the projector will provide a disk of light on a dark background. The slide used in the second projector is opaque except for a ringlike transparent area ½ in. on its inner diameter. This projector provides a ring of light that will just fit around the disk of light provided by the first projector. The projectors should be provided with means for varying the intensity of the projection, so that the experimenter can manipulate the absolute and relative intensities in the inner disk and outer ring at will.

When the intensity of the disk is held constant and the intensity of the ring is varied over a considerable range, the lightness of the disk may shift all the way from white to dark gray. If, instead, the intensity of the disk is manipulated and that of the ring held constant, the color of the disk shifts from light to dark.

It may be said, then, that the lightness of the disk, in general, depends upon that of the ring, and the lightness of the ring depends upon that of the disk.

When either the disk or the ring are presented alone, the mode of appearance of the light changes from that of a surface color to that of a luminous source. Opaque or surface colors apparently emerge only when two regions of different light intensity are adjacent, as in the disk-ring pattern, or when, for instance, two rectangular areas of light are contiguous along their long borders.

Color Contrast and Color-constancy. In the following arrangement you are provided with a demonstration of conditions pertaining both to color contrast and to color-constancy. A

narrow red ring is placed on a cardboard of medium gray. About 2 in. in front of this ring, and concentric with it, is placed an episcotister[1] whose solid sectors are bluish green. Thus, through the plane occupied by the episcotister, one can see the gray cardboard and also the ring. Of the four solid sectors of the episcotister each constitutes 50 deg. of the whole. This leaves four open sectors, each of 40 deg. It had been previously determined that a revolving disk with four 50-deg. sectors of bluish green and four 40-deg. sectors of red fused to provide a neutral gray (see Fig. 45).

From this, one might expect that the red ring seen through the revolving episcotister would appear gray. *What is actually seen, however, is an orange-red ring lying behind a transparent green film provided by the revolving episcotister.*

For the second part of the experiment (the contrast experiment), the same color elements are also provided. But in this case the revolving component is not the usual episcotister but a gray color wheel with a red ring mounted on it. Cemented on top of these elements is a sectored disk of bluish green comparable to the episcotister in the first arrangement. Thus what is provided in this experiment is a color wheel whose outer area (1) is gray, and the next area (2) is composed of alternate bluish-green and gray sectors. Within this is a narrow red ring (3) alternating with the bluish green of the sectored disk, and within this, of course, is the alternation of gray and bluish green again (4). In the very center (5) all is bluish green (see Fig. 45, left).

In this setup the retinal image colors and timing are the same as in the constancy experiment. There, however, the red ring was seen to lie *behind* the episcotister that formed a transparent green film. In the present contrast situation, all the elements are seen to lie in the same plane. Despite this difference, the color of the ring is the same in hue and saturation in both cases. This demonstrates that the presence or absence of the film has no effect on the perceived hues and saturations. That the colors

[1] An episcotister is a revolving disk composed of alternate blades and open sectors which alternately pass and obstruct light. In motion, the ordinary electric fan is, in principle, an episcotister. One can see objects beyond the fan through the path of the revolving blades.

in one case lie in separate planes and belong to the surfaces of two separate objects, and in the other case lie in the same plane as possible parts of a common pattern, does not alter the outcome. This has been used by some to demonstrate that the laws of color contrast and color-constancy are the same. This one demonstration may not, however, be sufficient to indicate that this is the case in all situations.

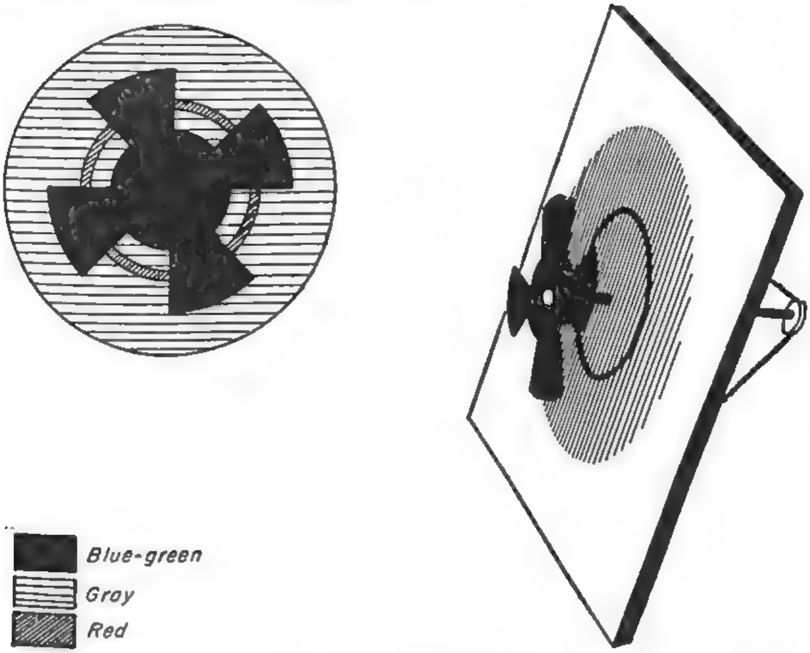


FIG. 45. Arrangements to show color-constancy and color contrast. The right-hand setup is a revolving disk of blue-green with open sectors. While in motion the disk presents the appearance of a translucent green *film* through which the gray disk (background) can be seen. On the gray background is a red ring also visible through the film. The left-hand diagram is a disk patterned to match the arrangement just described. Here, however, all the components of the field lie on a single plane and are not seen as separate objects. (*After Wallach and Galloway*, 1946. J. *exp. Psychol.*)

You have been given a few of the many phenomena of lightness- and color-constancies. The ones used represent some of the principles involved. Naturally no complete coverage of the topic could be made in this short chapter. Not being familiar with enough kinds of phenomena to cover all the possibilities of function, you are not able to make many dependable deductions

of your own. What has been given you ought, nevertheless, to excite your curiosity to know more about this very fundamental subject.

The problems that confront the investigators in the field of perceptual constancy have to do with making generalizations that fully but accurately describe the principles involved. Workers must first determine whether the field has been thoroughly enough canvassed to provide examples of all the principles that actually operate in determining constancy and its approximations before they can proceed with much assurance to formulate laws. Experimental psychology is in the midst of this procedure right now.

QUESTIONS

1. What is meant by "constancy" in perception?
2. What very general process is the basis of constancy?
3. In what general ways is constancy manifested?
4. What is the difference between lightness and brightness?
5. In what two ways may an object function with reference to light?
6. Name a common mode or way in which color is perceived.
7. Describe the experiment in which film color was involved.
8. In general, what does one do to minimize or obliterate lightness-constancy or color-constancy?
9. How are revolving disks used to manipulate color or lightness?
10. Would an artist, using color, want to be more given to perceiving constancies or less given to it?

CHAPTER 17

SIZE- AND SHAPE-CONSTANCIES

In the preceding chapter lightness- and color constancies were dealt with. They were considered as aspects of perception in which the individual reacts to stimuli as if they were things. It was pointed out that in order for there to be an experiential world of things, the organism has to utilize stimulation in such ways as to provide certain perceptual continuities and uniformities.

In this chapter we are about to give attention to two more uniformities necessary in the nature of things. They are *size- and shape-constancies*. There is a different kind of latitude involved in these two kinds of constancy. Objects may, in some cases, vary in size and still retain identity, but certain shape distortions are not tolerated without change from one *kind* of object to another or from one specimen to another. Objects may, therefore, tend to retain their shape despite the variations that occur in the images produced on the retina.

In an earlier chapter we used the example of familiar objects retaining their apparent size despite changes in size of their retinal images produced by substituting giant or miniature replicas. In such cases differences in size of the retinal image were utilized in seeing the object as the *same* size as before, but placed at different distances from the eye.

Whereas this experiment was performed in a darkroom where no assisting cues were obtainable, other experiments on size-constancy can be and are performed in situations containing various numbers and kinds of cues. For example, if one is asked to compare a 4 by 4-in. cardboard held in the hand with a number of cards fastened against a wall 15 ft. away, he can pick out a card that is the same size as the one he holds in his hand. He will not be far wrong unless there is something the matter with

his vision. It is obvious that to call a card that is 15 ft. away equal in apparent size to one held 2 ft. away is to compare two retinal images roughly 1 to 7.5 in linear dimensions.

Still another example of how perception operates was noted in observations made with binoculars. In such observations cubes do not look like cubes, but become flattened. The angles of the edges toward the observer are greater than 90 deg., while the other visible edges become less than 90 deg. Thus the "seen objects" are not cubes, but are other reasonable, solid geometric figures.

On the other hand, clock towers and cubical boxes with lids on retain their rectilinear features. In so doing, they must lose certain other features of shape to compensate for the retention of their full third-dimensional value. The principles determining which alternative of distortion will actually occur are quite complex and are not well enough understood to be reduced to verbal form.

A search for the factors that underlie size-constancy has been going on for a number of years. Various investigators have studied special aspects of the general problem. A number of them have studied the role of proprioceptive cues, such as possibly provided by the eye muscles in accommodation and convergence. Even eye-movements have been considered. One investigator called attention to the fact that an object (square) at a fixed distance from the eye differs in size depending upon whether one fixates on it, beyond it, or in front of it. With the far-fixation the square appears larger than with fixation upon it. With near-fixation the square appears smaller than with fixation upon it. The amount of convergence is taken by some to explain this phenomenon. Three-dimensional forms are said to manifest higher constancy than two-dimensional forms. Often higher constancy is found when using near targets[1] than far ones. Certain objects of standardized size in everyday experience tend to have a very high degree of size-constancy.

[1] Target is a term used to label what is looked at. It has various synonyms such as test-object. It is used especially when the user wishes to avoid specifying that what is looked at is seen as a definite object rather than a mere differentiation in the visual field.

Locke's Experiment. The following experiment will serve to represent one type of recent investigation of size-constancy. The targets used were gray cardboard squares standing erect at right angles to the line of regard. They were held upright by blocks fastened to the back. The standard square was a small one 15 cm. on a side. The other square was 16.5 cm. These were shifted from side to side at irregular intervals so as to obviate any asymmetry in the behavior of the observer. The squares were placed on a table 489 cm. long and 50 cm. wide and of the usual height (80 cm.). The observers sat 150 cm. from the end of the table. The standard square was placed 50 cm. from the end of the table next to the observers (200 cm. from the eye). The table surface was painted flat black. The visual field surrounding the table surface was provided by gray construction-board panels. These were placed on three sides of the table and reached almost to the ceiling. Illumination was provided by three 150-watt incandescent lamps 1 m. apart. The arrangement was such as to preclude shadows falling on the table. A white cloth screen prevented observation of the table and squares between trials.

Four different conditions were used. They involved the same procedure, however. The subjects, who were college girls, were instructed to report which of the two squares appeared to be larger. In so doing, they were not to answer as to which one they knew to be larger, or thought should be larger, but the one that actually looked larger. After instructions, the screen was drawn for a 3-sec. exposure of the squares. The observer immediately reported. The comparison square was moved backward from the standard square 5 cm. per trial until the observer consistently reported the physically smaller square, which remained in fixed position, as the larger.

Meter sticks attached to the edge of the table indicated the actual positions of the comparison square. The procedure, it will be noted, utilized the method of limits and involved four ascending and four descending series.

In condition A, the observers saw the table and the squares without any other forms or objects in the immediate field. In condition B, two posts were in the field. These posts were

25 cm. high and 2 cm. in cross section. They were placed to the side of the standard square and in the same plane. By intent they were to serve as reference points. In condition C, a 150-watt lamp illuminated the table from a position external to the area enclosed by the panels. The original three lights were turned off. By this means the light fell on the large comparison square, while the smaller standard square remained in shadow. In series D, the standard square was in the light and the larger comparison square was in the shadow.

For the 17 observers, the average distance at which the larger and more distant square became just smaller than the smaller standard square ranged from 252.5 to 367.5 cm. The standard square that was 200 cm. from the eye formed a tangent at the eye of approximately 0.075. ($15 \div 200 = 0.075$). If we divide the height of the larger square, 16.5 cm., by the average distance at which it was seen as just smaller, we have 0.051, or the approximate tangent which the comparison object forms at the eye ($16.5 \div 323.7 = 0.051$) for condition A.

Since the two squares were not at the same distance from the eye when they became just different in size, it is appropriate to see what the tangent of the standard square would be if removed to the distance of the comparison square. This means we consider it taken from its actual position at 200 to 323.7 cm. Then we divide 15 cm., its height, by 323.7 cm. and find we get the tangent of 0.046. This represents a smaller visual angle than the tangent of the comparison square (0.051).

For condition B, the average distance of the comparison square, when judged just smaller, was 354.3 cm. The tangent was 0.047. The tangent of the standard square projected to the distance of the comparison square was 0.042.

The results for conditions C and D were also computed. There was found to be a statistically reliable difference between the results obtained:

1. For the clear field and the field with posts.
2. For the clear field and the front-shaded field.
3. For the clear field and the rear-shaded field.
4. For the rear-shaded field and the front-shaded field.
5. For the rear-shaded field and the field with posts.

The obtained differences between the field with posts and the front-shaded field were not reliable.

The matter may be summed up as follows. A moderately high degree of size-constancy was found in all four conditions tested. The presence of posts (reference points) in the visual field increased the degree of size-constancy. When the field was differentially illuminated, the degree of size-constancy was increased in one set of conditions and decreased in the other. When the comparison square, but not the standard square, was illuminated, the constancy was increased. When conditions were reversed, size-constancy was reduced below that found in the clear-field situation.

Shape-constancy. Shape-constancy may be well illustrated by some experiments of Thouless in which he used a square and a circle cut out of cardboard and viewed from several different oblique angles.

The circle was about 39.75 cm. in diameter, and the square was 38 cm. In one experiment the circle was used and in another the square. These were laid on a black-surfaced table. The square had one corner directed toward the observer. The eyes were 48.5 cm. above the table. Three different positions were used. In situation A, the near edge of the test-object was 54.5 cm. from a point on the table just below the eye. In B, the near edge was 109 cm., and in C, it was 163.5 cm. Figure 46 indicates these positions.

Thouless called the circle and the square "real" objects. It is obvious that when they were viewed at an oblique angle, the shape of the retinal image of the circle was an ellipse, and the square was a diamond or lozenge. These shapes he called the "stimulus" objects. The objects the observer indicated by drawings were called the "phenomenal" objects.

The aim of the experiments was to determine the actual shapes that the observers saw. These, as was just said, were indicated by drawings. In Fig. 46 the phenomenal objects are solid black. From the illustrations it is evident that the phenomenal objects compared neither to the real objects nor to the stimulus objects (retinal images). In every case they lay somewhere in between the two in shape. Hypothetically the results could have been

either of the two extremes, namely, representative of the stimulus object or representative of the real object. The actual results represent what Thouless calls *regression to the real object.* This regression varied in amount depending upon the obliqueness of

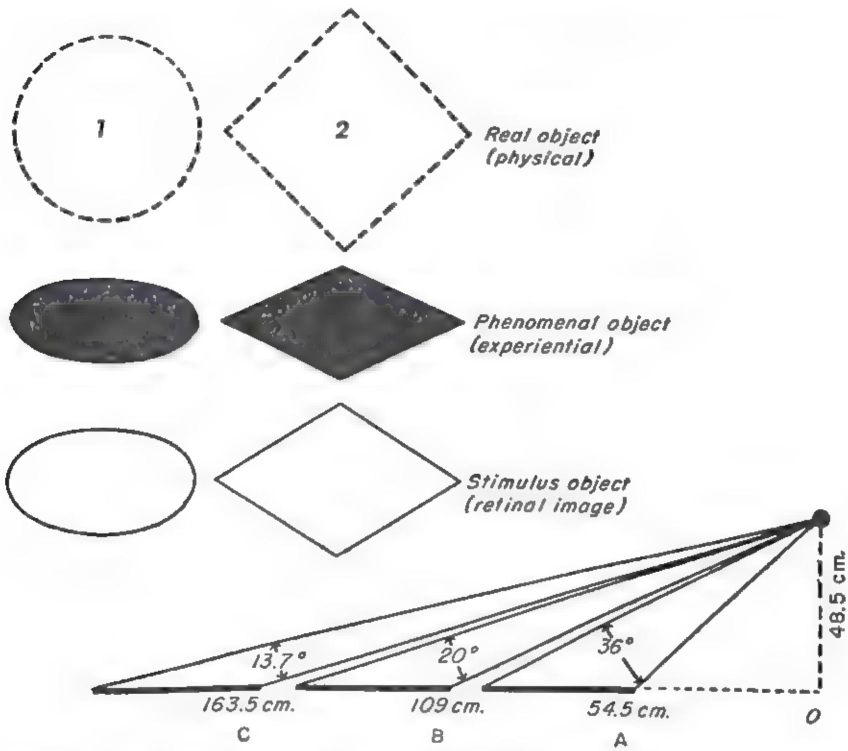


FIG. 46. Diagrams above (1) indicate the oblique orientation of a circle (disk) relative to the eye so as to produce an elliptical image on the retina and be seen as an ellipse, (2) indicate the same principle in the case of a square, causing it to be seen as a lozenge.

Diagram below shows the placement of the circle relative to the position of regard (eye position). The eye is above the table, and the circle is placed at the various distances indicated from the edge of the table. The same positioning was used for the square. (*Modified from Thouless*, 1931. *Brit. J. Psychol.*)

the angle of regard in viewing the real objects. The table below presents a summary of results on one subject. The first column, for each condition, represents the ratio of short to long dimensions of the phenomenal object. The second column represents the ratio of short to long dimensions in the stimulus object. The

third column for each condition indicates the *index of phenomenal regression*, a term that will presently be explained.

TABLE 12

Condition A			Condition B			Condition C		
P	*S*	*I*	*P*	*S*	*I*	*P*	*S*	*I*
Circle								
0.78	0.56	0.57	0.58	0.36	0.465	0.47	0.255	0.445
Square								
0.86	0.565	0.74	0.73	0.36	0.69	0.58	0.255	0.60

The Index of Phenomenal Regression. This term is the name for a quantitative way of expressing the relation of the shape seen by the observer to the shape of the real object. If S is the stimulus character, or the ratio of the short to long dimension of the stimulus object, and if P and R are the corresponding characters of the phenomenal and real objects, respectively, we have the following ratio or index:

$$\frac{P - S}{R - S}$$

If, for example, the real, the phenomenal, and the stimulus objects all had the same character, in the case of the square and circle the index would be:

$$\frac{1 - 1}{1 - 1} \text{ or } 0$$

If the circle was seen as a circle but the stimulus object was an ellipse with a character of 0.50, then the result would be:

$$\frac{1.0 - 0.50}{1.0 - 0.50} \text{ or } 1.0$$

But, in the actual case under condition A for the circle, the values were:

$$\frac{0.78 - 0.56}{1.0 - 0.56} = 0.50$$

Thouless decided, however, that the logarithms of the terms used would be more appropriate so that the simple index given above becomes

$$\frac{\log P - \log S}{\log R - \log S}$$

Using this form, the index of phenomenal regression turns out to possess the values indicated in the summary table already given.

Thouless made other experiments to check various questions in regard to the effects obtained in the one we have just described. Two of the main things we can say that were accomplished were the demonstration of the phenomenal object being somewhere between the real and the stimulus objects in the cases tested, and the formulation of standard quantitative indicator for expressing the actual character of the object as seen in relation to the so-called real object.

QUESTIONS

1. Describe a simple situation in which size-constancy is demonstrated.
2. Describe a simple situation in which shape-constancy is demonstrated.
3. Which kind or kinds of constancy are well manifested in optical magnification and which kind is often not so well manifested?
4. What is meant by "regression to the real object"?
5. What three kinds of "objects" did Thouless distinguish?
6. When the index of phenomenal regression is high, what may be said about constancy?
7. Could you apply Thouless' index of regression to the findings in Locke's experiment?
8. Was more constancy represented in the behavior of the circle or the square?
9. Suggest modifications of Thouless' experiment.
10. Which of the "objects" did the tangent of the standard square in Locke's experiment represent according to Thouless' classification?

CHAPTER 18

APPARENT VISUAL MOVEMENT

The experience of movement can be produced in many ways. Movement as an experience may pertain to the individual himself, to his surrounds, or to both, and is subject to many so-called illusions. The experience of bodily movement is a function of several sense-modalities, principally those of balance, vision, kinesthesis, and touch. Our concern here, however, is with the perception of movement of objects in the observer's surrounds.

Movement can be perceived either when something in the observer's surrounds is displaced or when nothing at all is displaced. Psychologists have called the experience in the first situation that of *real movement* and the experience in the second situation that of *apparent movement*. Only apparent movement will be dealt with in this chapter. Since vision is the chief modality for perceiving apparent movement of objects, our considerations will be limited to vision.

Apparent visual movement has been classified into essentially four forms: gamma, beta, delta, and alpha. Some of these forms may be studied by the serial presentation of geometrical figures (white on black, black on white, etc.), whereas others require areas of intense light emerging in a dark field such as provided by a darkroom.

Gamma movement is outward radial movement from a point (such as the center of fixation) when illumination is quickly raised, or when a perfectly dark field is suddenly illuminated. Gamma movement is also the movement in the reverse direction when illumination is lowered or terminated. Generally, the source of light in the visual field is a test-object or target of restricted area, but under everyday conditions the whole field may be illuminated. A room light may form the intensely

luminous area while the remaining field may be moderately bright.

Beta movement is the movement seen when two stationary areas or objects are successively exposed. These may be dark objects on a light ground or luminous areas in a dark field. When subtending an appropriate over-all visual angle, the proper timing of succession and the proper intensity (if luminous areas in a dark field), the two objects may be seen as a single object moving from one place to another. You are very well acquainted with examples of phenomena based upon such conditions. Electric signs and motion pictures form familiar everyday examples.

Delta movement is a combination of beta and gamma movements. Delta movement is the movement backward from the second of a successively presented pair of test-objects when the luminance of the second is greater than the first. Such a pair of objects will provide first for beta movement, as already described, followed immediately by movement in the reverse direction (in part, a kind of gamma movement).

Apparent movement is also producible by the successive presentation of two parts of a so-called optical illusion, such as the Müller-Lyer pattern shown in Fig. 47. You are familiar with the

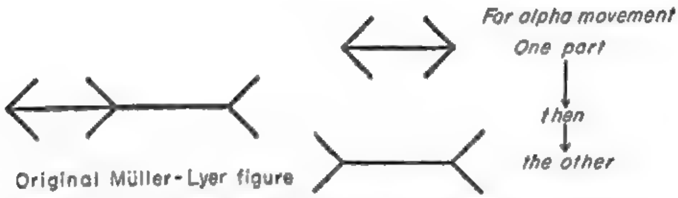


FIG. 47. An arrangement to produce alpha movement.

fact that the two parts of the horizontal line in the pattern, although geometrically equal, are seen to be unequal. If the pattern is divided as indicated, and shown successively, there will be an apparent lengthening or shortening of the line depending upon which component is shown first. Apparent movement under such conditions is called *alpha* movement.

Now we are ready to examine a specific investigation on the subject of apparent visual movement. Of these there are many,

but we shall concentrate upon an investigation carried out by DeSilva. It consisted in the study of a number of factors involved in the production of seen movement, not all of which we can deal with in detail. It is our purpose, rather, to find out how the investigation was conducted and what instrumentation was used.

DeSilva's Study. For his study, DeSilva employed a modified Dodge tachistoscope, a device for momentarily presenting two or more test-objects successively. The design of the Dodge tachistoscope relies upon the principle that a plate of glass is transparent when illuminated from behind and acts like a mirror when illuminated from in front.

Figure 48 is a diagram of the tachistoscope as seen from the top. It will be noted that there are three alleys that carry mirrors marked M_1, M_2, etc. An offset in each alley carries a slotted receiver for a card carrying a test-object. These are labeled S_1, S_2, and S_3. The glass plates, M_1, M_2, etc., allow the cards to be seen when illumination enters from the light sources B, B, B. But by arrangement, only one alley at a time is illuminated in this way. Hence, for example, when the near alley in the diagram is lighted and the others remain dark, glass plates M_1 and M_2 are transparent, and M_3 acts like a mirror. Therefore, an observer whose eyes are in position H can see the card at S_1, and not the cards at S_2 or S_3. If the middle alley is the only one illuminated, then M_4 and M_3 are transparent, and M_5 functions as a mirror enabling the observer to see what is on card S_2. The same principle holds for the third alley when it alone is illuminated. In each of the three cases the observer looks through two plates of glass and uses the third one as a mirror.

The ends of the alleys are provided with adjustable diaphragms whose openings are covered with frosted glass. The light entering the alleys from the sources B, B, B is diffused sufficiently by the glass and the matte-white side walls of the alleys to illuminate the cards at S_1, S_2, and S_3 quite uniformly.

A fall-shutter provides transient illumination of the three alleys in succession. This shutter is a plate of metal containing adjustable cardboard windows that pass by openings in the ends

of the alleys as the shutter falls. The durations of the light entering the alleys can be timed by narrowing or by widening the windows in the shutter. The shutter is guided by slots in the tachistoscope housing and is raised by a cord passing over a ball-bearing pulley. The shutter is held in the raised position until released by a trigger. The termination of the shutter's fall is broken and quieted by a padded spring platform.

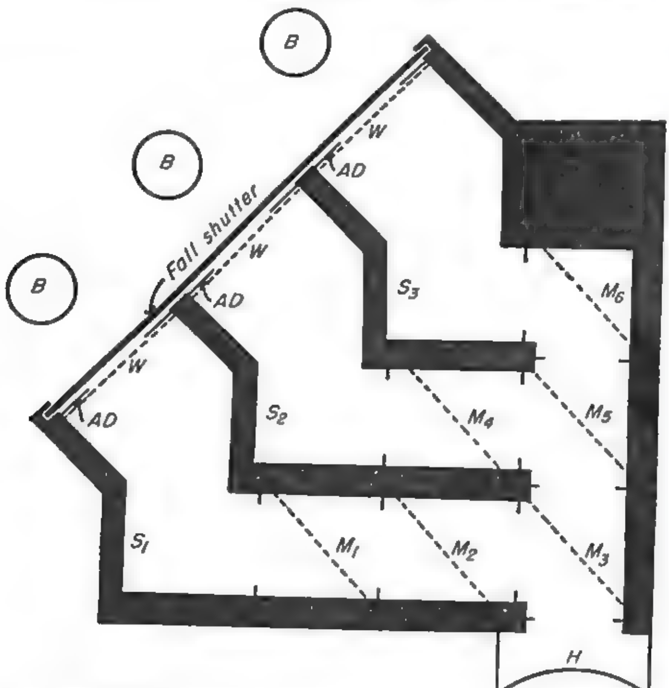


FIG. 48. A modified Dodge tachistoscope used by DeSilva. (For explanation of its operation see text.) (*DeSilva. By permission of Amer. J. Psychol.*, 1926, **37**, 471, *Fig.* 1.)

For DeSilva's study, the exposure times (durations of illumination) of the cards in alleys 1 and 3 bearing the test-objects were adjusted to 65 msec., whereas the duration of illumination of the black card in the second alley was 60 msec. The variation in the timing of the exposures was very little more than 2 msec. Cards at S_1, S_2, and S_3 were inserted through slots in the top of the tachistoscope.

Factors Studied. With the tachistoscope as described, a number of kinds of material was used for studying such factors

as the following: (1) The nature and development of the perception of movement; (2) the influence of repetition, fatigue, and form and meaningfulness of test-object; (3) influence of background; (4) the influence of position; (5) influence of fixation point; (6) influence of the form of the test-object on the direction of movement; and (7) the influence of voluntary predispositions on the direction of movement.

Procedure and Test-objects. To study the first-mentioned factor, that of the nature and development of movement, three test-objects were used, two of which are illustrated in Fig. 49. The first of these included two members, a card for alley 1, bearing the line *AA*, and a card for alley 3, bearing the line *BB*.

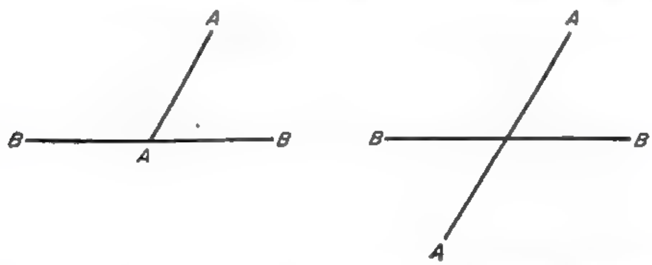


FIG. 49. The stimuli used in the first part of DeSilva's investigation of apparent visual movement. (*DeSilva. By permission of Amer. J. Psychol.*, 1926, **37**, 473, *Fig.* 2.)

Thus the sequence was from tilted radius *AA* to horizontal diameter *BB*.

The second test-object used was similar to the first, except that the line *AA* was the diameter of an imaginary circle instead of its radius. This test-object is shown in the right-hand half of the illustration.

The observers had to see several presentations of the sequence before they began to describe what they saw as unitary movement involving both members of the sequence. There was a gradual development of the "togetherness" of the two members into a unitary pattern of movement. The ends of the lines seemed to carry the movement. Several observers reported certainty of the motion, but were not sure of the order of appearance of the two members (tilted and horizontal lines), or the appearance of lines at all. (This is an example of the "pure phi-

phenomenon.") Some observers now and then reported movement in the direction opposite to that of the sequence of presentation. At first nearly all of the observers seemed to have set themselves to observe analytically rather than to take in the presentation as an over-all affair.

The second factor (or set of factors) involved a careful examination of repetition, fatigue, and form and meaningfulness of the test-object. For this purpose, eight different types of test-objects were used. These included the two used in the investigation just described. Among the others were a crossed diameter, a "windmill," and movement of a man saluting (see Fig. 50).

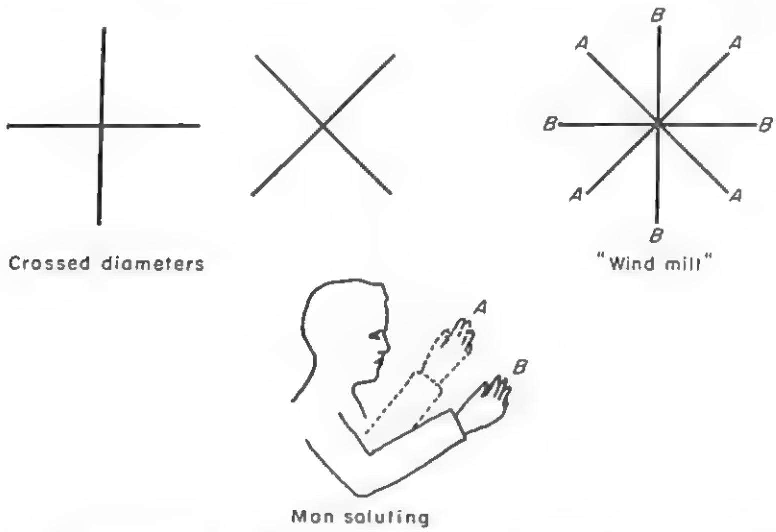


Fig. 50. The stimuli used in the second part of DeSilva's investigation of apparent visual movement. (*DeSilva. By permission of Amer. J. Psychol.*, 1926, **37**, 473, *Fig.* 2.)

Each test-object was presented for many trials, or until the observer reported unclear movement, after movement had once become clear. Reports of relative clearness were made at the end of every fifth trial on a scale from 1 to 5. One was "very clear movement"; 2 was "clear"; 3 was "moderately clear"; 4 was "vague"; and 5 was "unclear."

The period of greatest clearness in the perception of movement, in some cases, lasted only for 2 or 3 trials despite the attempt of the observer to maintain a constant attitude. In

other cases, however, maximum clearness persisted for as many as 60 to 70 trials. In many cases clearness increased again after it had begun to decline. The movement of a *radius* persisted as long as a *diameter*, but never was so clear. Movement of a radius permitted analysis into "succession" better than did movement of a diameter. It was more difficult to maintain a uniform passive attitude in observing a radius than in observing a diameter. The movement of two diameters persisted for more trials than the movement of a single diameter. The movement of a meaningful animate object, such as the arm of a man saluting, was more durable (lasted for more trials) than the motion of objects such as the radial and diametric lines. In this respect the durability of the movement of the windmill lay between that of the saluting arm and the radial and diametric lines (so-called *bare* lines). During early repetitions the movement of white lines on a black background was reported as relatively slow. After some repetition it was reported as faster. As the observer became more fatigued from observing, it not only became faster but less definite. With fatigue, in black on white or black on gray, the black seems to blur into the background, but this does not seem to happen with white lines on black, even when movement itself becomes less clear.

The third factor to be investigated was that of background. We have just mentioned the effects of white lines on black background vs. black lines on white or gray background, but only as they relate to durability.

In the investigation of the effect of background on the experience of movement, five different test-objects were used, some of which were the same as for the investigation of the preceding factor.

Movement of white on a black background is clearer and can be followed better than movement of black on white. The movement of white on black tends to be more impressive and attention-compelling than movement of black on white. The movement of black on gray is not so clear and easy to follow as black on white, although the difference is slight. On gray, the movement is reported as appearing to be faster than on white. The indistinctness of the arc of movement is more obvious on a gray

background than on a black background. In general, when a moving object stands out distinctly from its background, its movement is judged as slower than when it contrasts less with the background.

Injecting specific meaning into a line, such as seeing it as a kind of thread or cord, results in diminishing its speed of movement and renders its movement clearer and easier to follow. There is more of a "realness" to the movement of the thread or cord than to movement of a "bare" line.

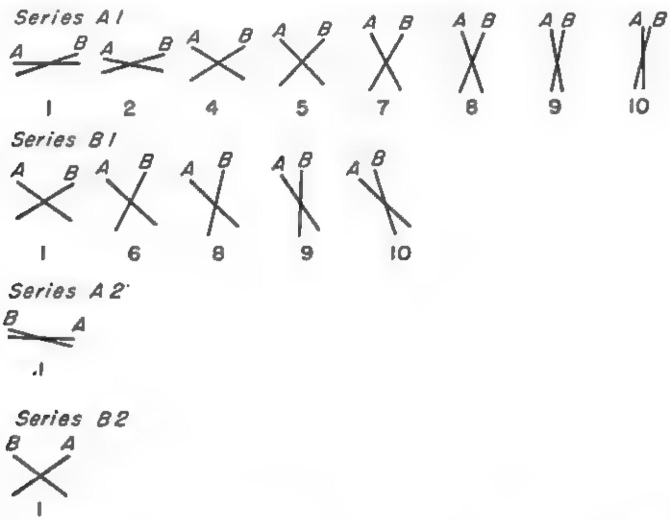


FIG. 51. Examples of stimuli used by DeSilva to determine direction of movement as dependent upon position. For example, in one of series $A1$, movement is counterclockwise from A to B. By the time stimuli 7, 8, etc. are reached the movement may change to being clockwise from A to B. In series $A2$, for example, the movement is clockwise from A to B, but further along in the series (not shown) the movement may begin to be counterclockwise. (*DeSilva. By permission of Amer. J. Psychol.*, 1926, **37**, 488, *Fig.* 5.)

Another factor studied was that of the position of the movement in the visual field relative to the point of regard. For this study a series of test-objects, as shown in Fig. 51, was used. The observers fixated a point representing the mid-points on the diameters AA and BB. With the series of test-objects, such as indicated in the illustration, the question of whether movement could be perceived better when it originated in the horizontal or in the vertical axis could be answered. The test-

objects also provided for determining whether movement to the left or to the right was easier to perceive; *i.e.*, whether successions of presentations toward the left or toward the right tended to institute clearer movement.

Movement in series $A1$ and $B1$ was counterclockwise, and movement in $A2$ and $B2$ was clockwise. Movement produced by test-object arrangements from 1 to 5 inclusive was called *lateral*, whereas movement in 6 to 10 was called *transverse*.

The broad outcome was that the compellingness of the presentations toward seeing movement was not the same in every position in the field. There was a decidedly greater tendency for lateral movement than for transverse movement to occur. The *direction* of seen movement tended to be reversed in keeping with whether the initial line (AA) appeared tilted slightly to the right or slightly to the left of the vertical. It seemed as though a line slightly off vertical was expected to fall rather than to rise. The direction of movement, therefore, depended upon which side of the vertical the line originated. In diametric movement the tendency was for the attention to be centered upon the area above the axis of rotation, and it was this part of the line that tended to fall when AA–BB line sequence was presented.

The influence of position in the visual field upon the ease with which apparent movement could be initiated was tested also by shifting the *fixation point*. The test-object for this was

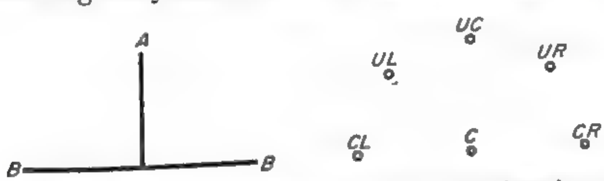


FIG. 52. Test-object (left) for studying the effect of location of fixation point or movement. The various fixation points used (right). C is the fixation when the observer looks at the intersection of vertical line A and horizontal line BB. (*DeSilva. By permission of Amer. J. Psychol.*, 1926, **37**, 473, *Fig.* 2.)

an arrangement (see Fig. 52) in which A (the first line presented) was a vertical radius of a circle having a diameter the length of BB, the second line to be presented. Six different fixation points were used as indicated in the figure.

The location of the fixation point determined very significantly the over-all experience of movement. The most favorable area of fixation for perceiving movement as a whole was the region in which the completion of the movement occurred. The next most advantageous area was the center of rotation of the line (point *C*).

Conclusions. DeSilva divided his conclusions into those regarding the determinants of movement "not in the stimulus," and those determinants within the stimulus. In the first group, four items were included and in the second group, eight items were listed. Among the nonstimulus factors were analytical vs. synthetic attitude, passive vs. active-striving attitude, and inertia of the organism against carrying movement across the vertical axis when the members were radially moving lines. Among the stimulus determinants were the artificiality of the tachistoscopic technique, complexity of the vehicle of movement (whether lines or something else), repetition in the presentation of the test-object, the location of fixation points, and the nature of the background.

QUESTIONS

1. Distinguish between real and apparent movement.
2. Distinguish between the four kinds of apparent movement.
3. Explain the basic principle upon which the Dodge tachistoscope operates.
4. Why were there *three* parallel alleys in the tachistoscope used by DeSilva?
5. What is meant by the "pure phi-phenomenon"?
6. What is a fall-shutter?
7. What could be substituted for the fall-shutter in DeSilva's apparatus?
8. What was the observer able to do in DeSilva's experiment to enhance the "realness" of the apparent movement of a line?
9. What form of apparent movement was DeSilva mainly dealing with?
10. What was the most favorable location of fixation for perceiving movement as a whole in DeSilva's experiments?

CHAPTER 19

THE REVERSAL OF APPARENT MOVEMENT

Gamma movement was one of the forms of apparent movement defined in the preceding chapter. Gamma movement occurs as a figure emerges upon a ground. In fact some investigators have used the conditions that produce gamma movement as a means for studying certain figure-ground phenomena. Some workers believed the radial expansion effect in gamma movement to be an inherent feature of figure emergence. Newman was one such worker. He varied the time during which the light area emerging as a figure increased in intensity. With his experimental arrangement, instead of the figure emerging by virtue of a sudden presentation of light at full intensity, the transition to full intensity required as much as 1.8 sec. This gradual rise of intensity was provided by a disk with a wedgelike open sector that passed a broad beam of light. As the disk revolved, the sector allowed an increasing portion of the beam to reach the test surface forming the figure. The use of this arrangement disclosed three stages in gamma movement: (1) The subjective coming on of the luminance; (2) stage of growth in which the lighted area (figure) subjectively expands to its full size; and (3) the apparent motion of the figure area toward the observer. Increasing the ultimate intensity of the figure area enhances this movement, called depth movement. Intensity transitions within the limits of 200 to 800 msec. are optimal for depth movement.

One of the experiments used by Newman to demonstrate the connection between the direction of movement and the emergence of figure was one in which two rectangles of light were made to appear upon a dark field. Figure 53 pictures the spatial and intensity arrangements. When the rectangles were illuminated, there was a complex pattern of apparent movement in the field.

One aspect of this was the broad movement outward in all directions over the field. Another aspect was the movement within the restricted areas around the rectangles. This movement proceeded in all directions locally from the two rectangles. Since there were two rectangles, some of their movement was in opposing directions. What the observer reported was dependent upon what area he fixed his attention.

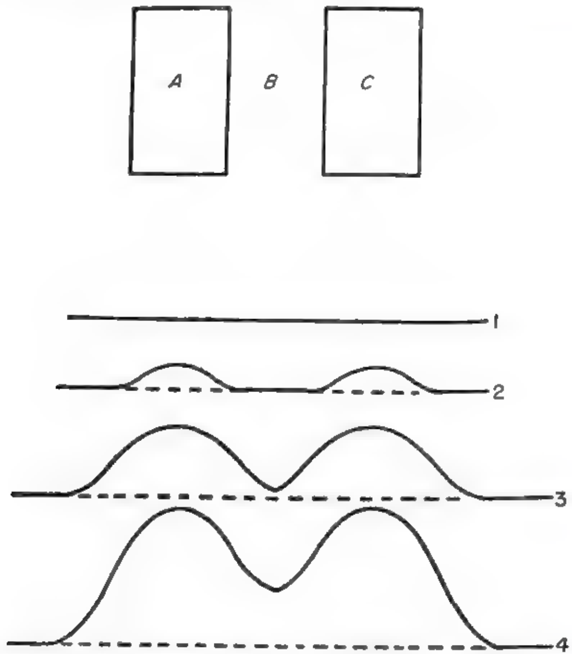


FIG. 53. Diagram to show arrangement of target in which two bright rectangles, *A* and *C*, are made to emerge on a dark ground. The space *B* between them may also be seen as a rectangle. Items 1, 2, 3, and 4 indicate the stages of spread of light on the retina as the rectangles emerge. (See text for further discussion.)

If the dark area between the two light rectangles is also seen as a rectangle (a figure), the observer can see the area expand as it comes into existence. It is usually seen to contract, however, as it emerges, even when it is seen as a rectangle. This showed that it was not compulsory for a figure to expand as it emerged.

To understand the literal stimulus situation, it is necessary to consider the sequence of illumination stages on the retina as the two lighted rectangles emerge. Figure 53 is intended to portray

this. Areas A and C are the rectangles that are lighted up momentarily. Area B, between A and C, can be seen as the dark rectangle, already mentioned, when A and C are lighted. Let diagrams 1, 2, 3, and 4 indicate the sequence of events as the areas A and C are rising to full intensity. Diagram 1 represents the prelighted situation. In diagram 2, the first emergence of light on the retina appears in a and c, the centers of the areas that will represent A and C. In diagram 3, the areas are becoming more intense and suprathreshold light is spreading out to cover more and more territory. Light is spreading to reach b, the center of rectangle B. In diagram 4, the light of both image areas has expanded to reach b and, in overlapping, has risen to an appreciable intensity. This series of diagrams shows the fact that, in becoming increasingly intense, light has, in effect, traveled across a portion of the retina. It has traveled outward in all directions from the center of the retinal areas that represent A and C, and in so doing has provided a sequence of illumination changes across the area between A and C. It is, therefore, not difficult to understand why movement of some sort or other is seen in that area.

The organism is well able to utilize these retinal events to see either contraction or expansion of B as areas A and C light up. Whereas the spatial sequence in the rise of light intensity from a and c, just mentioned, had to do with area B, there was a radial spread of illumination in all other directions giving rise to the larger aspects of the gamma effect. Stray light within the eye participates in carrying the movement somewhat beyond the nearby regions around the image of the lighted figure.

Another type of crucial experiment that was performed by several workers is what is known as a Lochversuch (a hole experiment). A Lochversuch represents the following set of conditions. In a wall or screen is placed a hole of the proper size and shape to form the desired test-object or target. The area of the wall or screen around the hole is a uniform white and is kept at the desired level of illumination. Beyond the wall or screen is a second surface that can be illuminated at any desired level. If the level of illumination of this second surface and the wall are equal, the hole is not visible, particularly if the edge of the hole

is sharp. If the illumination of the second surface is suddenly lowered or terminated, a black object formed by the hole looms up. It is obvious, then, that such an arrangement is a set of conditions providing for the emergence of a black figure on a white ground. Such a set of conditions then can be used to test whether a dark figure expands just as does the usual light figure. In the present case the sudden change is one of diminution or termination of light rather than its sudden onset or increase in intensity. In viewing the emergence of a black figure on a light ground, one sees it expand as it develops. In fact, it was possibly this sort of an experiment that played a part in forming the conviction that figures always expand (involve gamma movement as they emerge).

Gamma Movement with No Figure Involved. Another aspect of the figure-formation contention was that gamma movement was *dependent* upon the emergence of a figure. This was to say that where illumination changes occurred but no figure emerged or decayed, there could be no gamma movement. Bartley tested this assertion by using a very broad field—90 deg. or more in visual angle. This area was provided by an opal-glass screen uniformly illuminated from behind and lighted intermittently. In viewing this field, the observer saw it expand as the light came on and contract as the light went off. Since the field was unstructured, gamma movement in it showed that such movement was not dependent upon the emergence or decay of a figure upon a ground.

Reversal of Gamma Movement. The following is a description of an experiment on the reversal of gamma movement. Since to produce gamma movement in any case, there had to be a set of retinal (stimulus) conditions providing for a temporal sequence in initiation of optic nerve impulses representing adjacent regions. Bartley supposed the direction could be reversed by manipulating retinal illumination conditions in certain ways. A series of conditions, therefore, representing the essential varieties of retinal pattern set up by illumination were tried out. One of these illumination conditions was a homogeneous field of light provided by a large opal-glass screen. Another was a pair of small lighted disks, one of which was fixated and

the other imaged about 15 deg. from the fovea. A third condition was provided by a series of disks of increasingly great diameters. A fourth condition was provided by use of a large disk of light, the intensity of which was tapered so as to be less at the center than at the edge. Several other conditions were also tested, but these need not be described here.

It was recognized that even with a uniform field of light presented to the eye, the resulting retinal stimulation would not copy this uniformity. Light entering through the pupil obliquely is not so effective as that which enters more nearly normally (Stiles-Crawford effect). In addition to this, a retinal gradient in the receptor population from fovea to periphery is presumed to exist, so that the periphery may be slower to react than more central regions. If these two factors operate, then uniformly distributed light in the external visual field would set up a *temporal sequence* of activity from fovea to periphery. This would eventuate in part from the tapered retinal stimulation, and in part from the inherent retinal latency gradient from fovea to periphery. Apparent movement would be expected from such conditions.

The apparatus for producing the illumination or target patterns was the following. A lamp house with a circular, opal-glass window 2½ in. in diameter was the source of the light falling on a plano-convex lens 3 in. in diameter. A second lens of the same kind was located ½ in. from the first. Between the two a disk with a single open sector was placed. It slowly revolved, and as the open sector appeared in front of the lens, a short but gradual onset of light was produced. Thus the light tapered upward in intensity rather than coming on extremely abruptly. At a distance from the second lens a circular image of light fell on an opal-glass plate. A second plate of the same kind was located a few inches still farther away. The lenses, the opal-glass plates, and the path of the light were housed so as not to light up the room unduly. Figure 54 shows the optical scheme. At the third glass plate the housing was expanded so as to accommodate a fourth opal-glass plate 12½ in. square. The face of this plate was masked by a diaphragm with a circular opening 11 in. in diameter. This area (viewing screen) was large enough to

fill a considerable portion of the visual field when an observer was 12 in. or more away from it. The use of the opal-glass plates 2 and 3 and the lenses obviated the "travel" of a pattern of uneven illumination that would be seen on the opal-glass viewing screen as the edge of the open sector of the disk passed across in front of opal-glass plate 1. With the present arrangement, the only variation produced on the screen was one of either increasing or decreasing intensity.

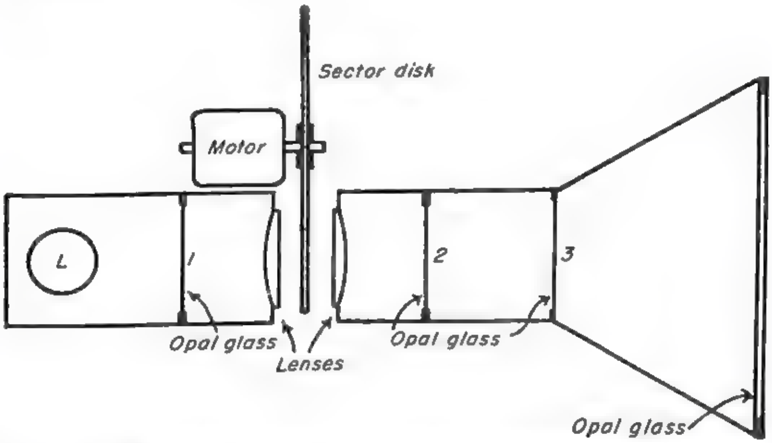


FIG. 54. A long-section diagram of Bartley's apparatus to provide a broad field of illumination that can be made to rise or fall gradually in intensity.

Condition 1 resulted in the production of the usual expanding gamma movement at the onset of illumination, even though the illumination source was homogeneous and occupied an extensive portion of the visual field. It will be remembered that according to the already described nature of the retina, gamma movement was to be expected, and in the direction observed.

Condition 2, in which the observer looked directly at one disk as it appeared and saw the other one 15 deg. to the side, produced a very significant result. The two equally intense disks were presented *simultaneously*. They were produced by two openings in a cardboard placed over the large opal-glass screen in the apparatus that we have described. It should be apparent that the setup in this case provides a good test for whether or not the peripheral retina lags behind the fovea (central retina). The disk fixated did loom up earlier than the other one and manifested

little movement. Thus lag definitely was shown. The movement in the second (peripheral) disk was marked. It spread out very definitely from the disk, forming a halo around it. If the behavior in the various portions of the field was mainly dependent upon *figure formation*, the two figures ought to have acted quite alike. As it was, the result seemed to indicate that what happens in connection with a figure is largely dependent upon the part of the retina involved.

Condition 3, in which disks (figures) of various sizes were produced, is complementary to condition 2. The purpose of using disks of various sizes was to produce images covering different amounts of the whole retina and to place the borders of the disks at different points along a possible center-to-periphery gradient of the retina. If the retina is not equally sensitive from center to periphery, the results of this procedure ought to provide useful information regarding the matter. It is known that small, intense stimuli produce figures that appear to expand when they emerge. That is, the borders of such figures move outward as the figures rise to full intensity. This description is so often given regarding gamma movement and figure formation that it has come to be taken as axiomatic for all conditions. Were we to use disks whose borders lie far out toward the periphery of the visual field, it is possible that the phenomenal characteristics of the emerging figures would be decidedly different. Thus to find this out and to encounter any other unthought-of phenomena, a series of disks including a variety of diameters from small to large was used. This series of disks demonstrated that figures could be so large as to contain much of the gamma movement *within* their borders. The usual *expansion outward of the border* of the figure as it emerges does not occur when the figure is very large. What movement there is occurs *within* the figure. Hence the prevalent idea that a figure must expand as it emerges does not hold. This expansion is a phenomenon quite largely dependent upon the region of the retina involved. The prevalent idea that border expansion must occur as a figure develops must be replaced by the idea that the behavior of the figure depends upon how much of the retina is involved and particularly what part of the retina is involved at the figure's contour.

Condition 4, as you will recall, consisted of the use of a very large field. For the purpose at hand, it covered about 44 deg. of visual angle. It was formed by placing a number of layers of tissue paper, each one somewhat smaller in diameter than the previous one, over the opal-glass screen already described. This formed a light barrier that was progressively greater from periphery to center. Accordingly, when the lighted field was presented to the eye, it was tapered upward in intensity toward the periphery. When this field was viewed, gamma movement, instead of being a radial expansion, was definitely a centripetal contraction as the light came on. Thus, by manipulating the relative intensities of the center-to-periphery gradient, gamma movement was reversed. This result was a demonstration that outward expansion is not a compulsory product of raising light intensity, but is dependent upon how the various parts of the retina are involved as a stimulus surface.

The fact that gamma movement can be reversed by appropriate conditions leads to the supposition that, with the proper tapering of the illumination gradient on the retina as a whole, the now ubiquitous phenomenon of gamma movement could be obliterated as a consequence of turning light on and off.

QUESTIONS

1. What three stages in gamma movement did Newman disclose?
2. What is meant by "figure ground"?
3. Does gamma movement occur when no figure is involved?
4. What retinal factors does Bartley declare gamma movement is based upon?
5. What are the conditions necessary for the reversal of the direction of gamma movement?
6. What is the significance of being able to reverse gamma movement?
7. What principle was indicated when two equal-sized and equally intense targets were used to produce gamma movement, and one target was fixated?
8. Where does gamma movement occur when large areas are used as targets?
9. What bearing have Bartley's experiments on the assertion that gamma movement is necessary in figure formation?
10. What did the Lochversuch demonstrate?

CHAPTER 20

THE RELATION BETWEEN SPACE AND TIME IN PERCEPTION

In the two previous chapters we reviewed some of the facts regarding the perception of movement, particularly some of the unusual conditions involved in its production. In the present chapter we again deal with the relation between space and time, although in a somewhat more direct way. In one way or another, perceptual reactions are always dependent upon the operation of space-time factors.

In the present chapter we shift our scene of interest from vision to the sense of touch. In this modality we can very clearly demonstrate space-time relations. For this purpose, we go to the investigations of Helson and King, who studied a phenomenon that Helson had previously called the *tau effect*. The effect was given its name to prevent it from being passed off as an illusion. We have already been dealing with many phenomena that the man-in-the-street would call illusions. The common implication is that illusions are phenomena that are not real, or at least not so real as other phenomena, or perhaps not real in the same way or sense of the word. Illusions are thus taken to be phenomena in a distinct functional class by themselves. If illusoriness is to be defined as a discrepancy between "knowledge" and sensation, then there are many degrees of illusoriness, and many of the phenomena that have never been classed as illusions partake of this discrepancy in some form or degree. For this reason, if for no other, the demarcation of some phenomena as illusions, and some as not, is based upon an incomplete knowledge of sensory phenomena.

Illusions are actually based upon the fact that conditions for perception are such that very definite conflicts occasionally arise

within a single sense-modality, or between sense-modalities. A thing feels big and looks little, looks crooked but "is" straight, etc. All this means that we are able to use more than one frame of reference in dealing with object-properties. One object-property we take to be real; the other less real, or artificial. The results which we believe to be artificial tend to be passed off as curiosities or entertaining stunts. We do not realize that to know the universe is to be aware of the conditions that make for perceptual conflicts. So-called illusions therefore have special scientific value.

We should be aware of the *kinds* of factors that underlie perception. Size, shape, intensity, and the like were tacitly assumed to be properties of perceptual situations, but time is often ignored as a factor that operates interchangeably with those of space and intensity in stimulation.

Helson and King put to precise test the following simple situation involving time as such a factor. If, on the forearm of an observer, you touch three points in a row, you thereby mark off two space intervals; the first lies between contact one and two, the second between two and three. If the observer has his eyes closed so as not to see his arm, you can touch these three points in such a temporal sequence as to have the second space interval much shorter than the first. Let us say that you quickly touch the first two points and wait a little before touching the third point. Then you ask the observer to tell you whether the second *space* interval is longer, shorter, or equal to the first one. The observer may tell you that the second space interval is longer than the first (even though spatially it is much the shorter). This phenomenon is called the tau effect and was investigated experimentally with precise timing provided by an electro-mechanical instrument called a kinohapt. A kinohapt, of the type used, is a device that carries three solenoids[1] whose position

[1] A solenoid is a coreless electromagnet. In place of the fixed core, a rod may move up and down in the core space. When the current is turned off, the rod falls, and when the current is again turned on, the magnet's field draws the rod up into position within the core space again. By this means a quick and precisely timed movement can be produced. This can be utilized for applying contact points to the skin as was the case in the present investigation.

in the three dimensions of space can be controlled. Each of the solenoids operates a blunt stylus for contacting the skin. The timing of the action of the solenoids, and thus the application of the contact-styli to the forearm, is controlled by a make-and-break circuit. In the investigations of Helson and King this circuit was activated by a phonograph turntable. The timing was adjusted by changing the positions of contacts on the turntable. The period of contact of the styli on the skin was about 100 msec. The weight of each stylus was about 11.5 g., and each was allowed to fall about 2 mm. The arm was held steady in an open plaster mold. A separate mold was made for each observer and conformed to the lower surface of the arm when the arm was in a horizontal position. In some experiments the ventral and in others the dorsal surface was used for exploration.

The first spatial interval we shall call S_1 and the second S_2. The judged spatial interval will be called S_j. The value for S_1 was 30 mm. Starting with $S_1 = S_2$, S_2 was lengthened or shortened 5 mm. per trial until it was judged correctly for 20 consecutive trials, the number of trials used for each value. There were six observers, one of them a blind student who had not seen since the age of ten. Table 13 indicates the form used for most of the data. It will be noted that one of three judgments was allowed the observers. They were always to judge the second *spatial* in-

TABLE 13

Name *S* intervals	Shorter	Equal	Longer
30–30			
30–35			
30–40			
30–45			
30–50			
30–55			
30–60			
30–65			
30–70			
30–75			
30–80			

terval with reference to the first. The change in spatial interval was continued step by step until the observer was able to make 20 consecutive correct judgments. This is designated as 100 per cent in the table. All the other data are also given in terms of percentage correctness.

Figures 55, 56, 57, and 58 (left) (pages 198–199) show curves that were drawn through points on the plots by inspection (by your author). Helson and King confined the data which they obtained to tabular form. The plots presented here are summary plots combining the reports for the six subjects. It will be noted that the time intervals between contacts are different for each graph. In the first graph they are 500 and 200 msec. In the second graph they are 500 and 250 msec.; in the third, 500 and 300 msec.; the fourth, 500 and 350 msec.; and in the fifth, 500 and 400 msec. In other words, the two time intervals started out by being quite different and in successive experiments became more nearly alike.

As was said earlier, the spatial value of the standard or first interval was 30 mm. In the figures the abscissas are plotted from 0 to 50 mm., meaning that the second spatial interval ranged from 30 to 80 mm. The abscissa designates only the absolute difference between the two spatial intervals.

It will be noted that since the second *temporal* interval was *less* than the first, the second *spatial* interval had to be *greater* than the first in order that all 20 judgments in a series might be correct.

Likewise, as the difference between the two temporal intervals became more nearly alike, the difference in the two spatial intervals needed for this correctness in interval judgment became less. Whereas in some experiments the second temporal interval (t_2) was made longer than the first (t_1), in other experiments it was made shorter. Figure 58 (right) shows a graph of results obtained when t_1 was 250 msec., and t_2 was 500 msec. It, like the preceding freehand sets of curves, provides a rough but useful indication of the general relationships between the percentages of correct judgments and the relation between time and space intervals. If we were to take the graphs literally, we should note a difference between the two sets, one for the $t_1(500)$ $t_2(250)$ and

the opposite, $t_1(250)$–$t_2(500)$. Such a difference might be expected. In experiments such as those of weight lifting, it makes a difference whether the heavier or lighter of the two weights is lifted first.

Were we to consider the investigation as an ideal and formal example of psychophysical procedures, certain refinements would be called for in order that thresholds, etc., could be strictly determined. Curves fitted to the data by procedures other than mere inspection would be required. We must not lose sight of the purpose of the investigation, namely, to demonstrate the reality of the phenomenon (tau effect) using quantitative procedures only to the extent necessary to accomplish this purpose. The investigation of some problems, were psychophysical methods to be carried out to their limits, would be almost too extended to undertake. This present phenomenon is probably such an example.

Table 14 is a summarizing table that will help in getting a comprehensive picture of the phenomenon from a little different angle.

TABLE 14

t_1/t_2		S_2/S_1	
500/200 or	2.5	75/30 or	2.5
500/250	2.0	65/30	2.17
500/300	1.66	55/30	1.83
500/350	1.43	50/30	1.66
500/400	1.25	45/30	1.50

If we call the tau effect the compensatory effect of time for space, then it will be seen from the table that it is much less when the ratio of S_2 to S_1 becomes small. Thus by looking down the appropriate columns in the table, we can see that when conditions are not so favorable for the tau effect, the spatial and temporal ratios become more widely discrepant from each other. For example, 2.50 and 2.50 match, but 1.25 and 1.50 are somewhat different.

The authors also applied their procedure to answering the question of what would happen if two tactile locations instead of three were used. The order of presentation was 1–2–1 instead of 1–2–3. The question is whether the tau effect will appear with

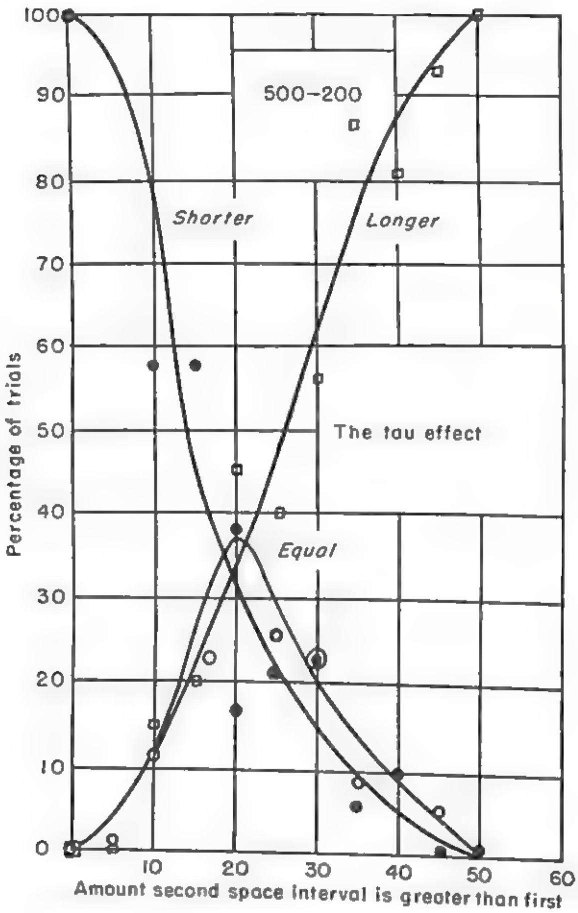


FIG. 55. Freehand curves drawn through points representing Helson's and King's data on the tau effect when the two time intervals between three tactile stimuli were 500 msec. and 200 msec. From the curves it will be seen that S_2 must be 50 mm. greater than S_1 before it is judged greater 100 per cent of the time. S_1 is 30 mm.

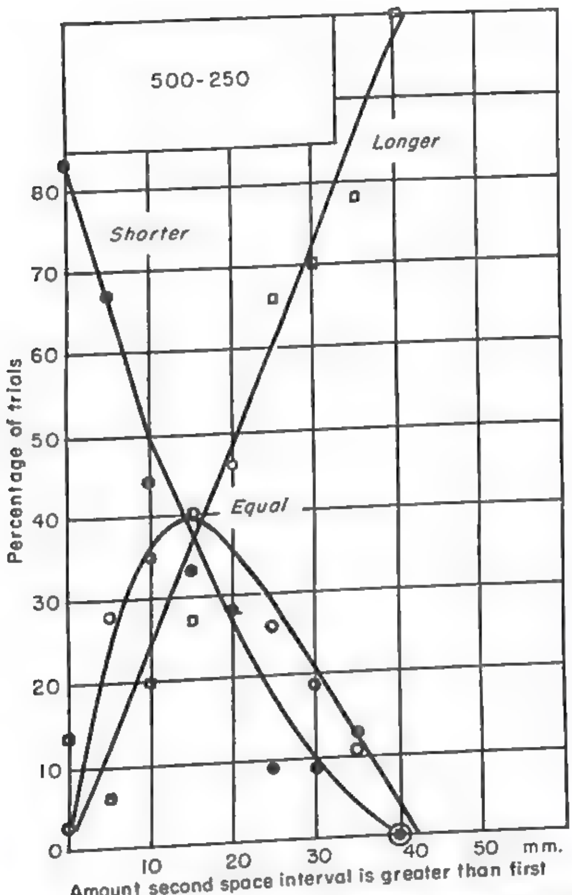


FIG. 56. Graphs to indicate the relation between space (S_2) intervals and percentages of judgments of equal, shorter, and longer, when the two time intervals are 500 msec. and 250 msec., respectively. The standard space interval is 30 mm. The curves are drawn freehand from inspection.

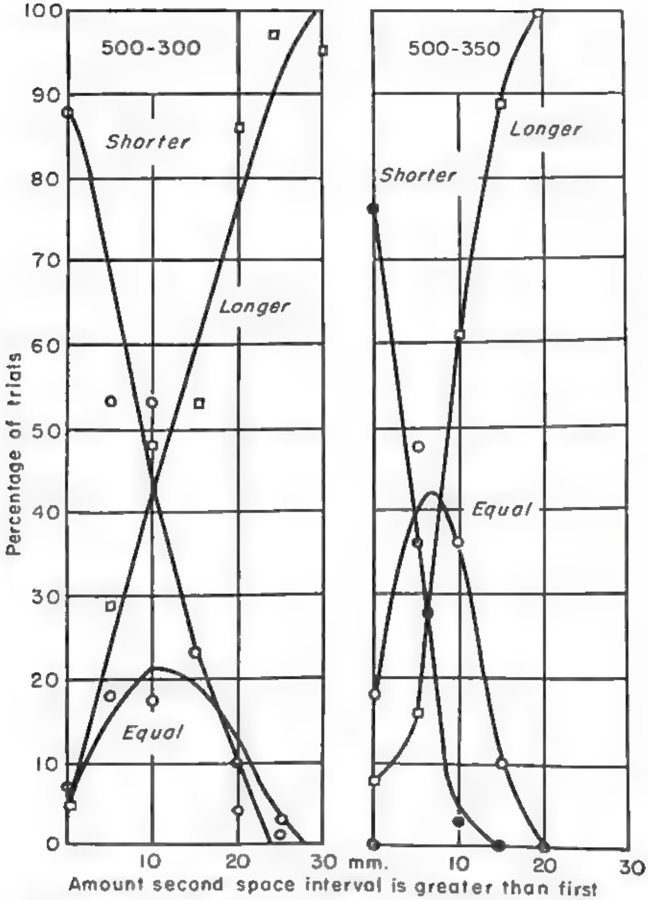


FIG. 57. Freehand graphs to show the relation between space intervals (S_2) and percentages of judgments of equal, shorter, and longer, when the two time intervals are 500 msec. and 300 msec., and 500 msec. and 350 msec.

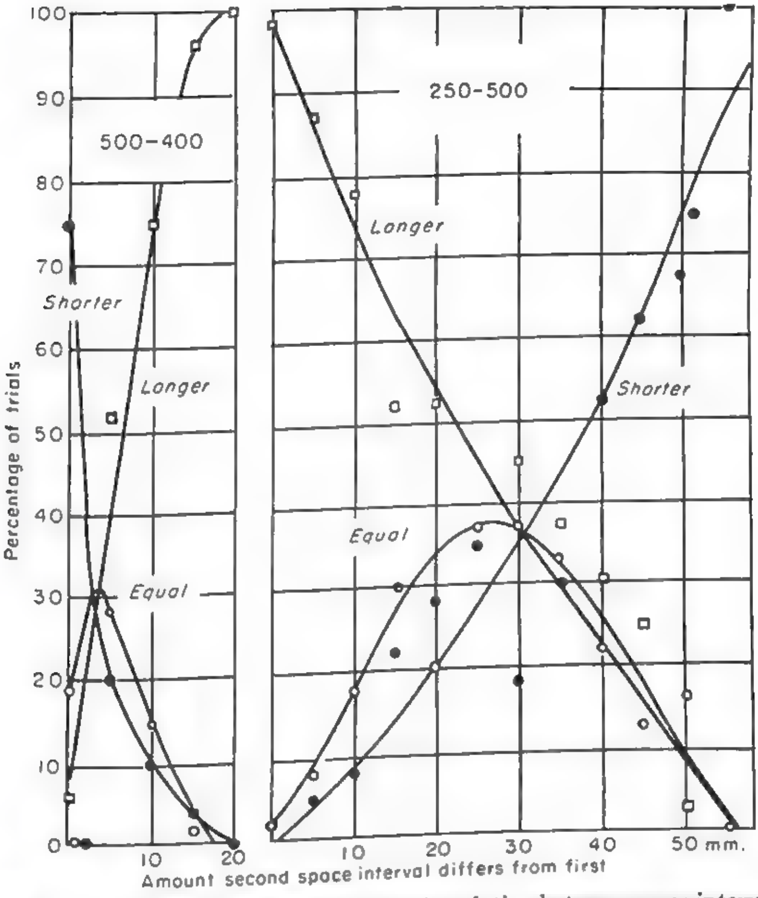


FIG. 58. Freehand graphs (left) to show the relation between space intervals (S_2) and judgments of equal, shorter, and longer, when the time intervals are 500 msec. and 400 msec. Graphs (right) to show the reversal of the tau effect when the second time interval is longer than the first (250 msec. and 500 msec.).

this procedure when the values of the two time intervals are made different. With this procedure they used 1,000 msec. for the first time interval (t_1) and 400 msec. for the second time interval (t_2) in one experiment. In another, $t_1 = 250$, and $t_2 = 100$ msec. The authors reasoned that if the tau effect did appear, then the doctrine of *local signs*, according to which every point on the skin has its own peculiar quality, should be abandoned.

The outcome indicated that, although the reports were more variable than when three points were used, the factor of time did influence it. When there was a difference in timing, the spatial interval 1–2 was not perceptually equal to 2–1, although geometrically it was the very same interval. Even when the judged spatial intervals were reported equal, the impression persisted, in many cases, that the point touched in the third contact was not the one touched in the first. This experiment constituted one of the cleverest and cleanest arguments on record against the old rigid theory of local signs.

As a final experiment, the authors rearranged the spatial configuration of the three points touched. Whereas they originally were in a straight line, in this experiment they were made to form the three corners of an equilateral triangle. Using this arrangement, the authors varied the timing of the intervals from equality to various amounts of inequality. Just so long as the time intervals were kept equal, the perceptual impression was that the three contact points formed the corners of an equilateral triangle. When the timing was shifted, then the relative lengths of three sides of the triangle became altered. The triangle ceased to be equilateral.

Here again (in this triangular arrangement) we have a cleverly conceived set of conditions for readily demonstrating the tau effect and some of its implications.

In interpreting the tau effect, you must remember that the activities underlying stimulation and the perception of touch are essentially neurophysiological. The matter of timing neural activities is crucial because the inputs from different peripheral areas (cutaneous, in this case) function in accordance with how they combine with inputs from elsewhere when they reach given

synapses. Inputs reaching synapses simultaneously are differently effective from those which reach them nonconcurrently.

QUESTIONS

1. What is the tau effect?
2. Can you describe a simple way of demonstrating, without apparatus, the tau effect to an observer?
3. What significance has the tau effect on the psychology of perception?
4. What is a kinohapt? What does it do?
5. What were the variables in Helson's and King's investigation?
6. Did Helson and King strictly follow any formal psychophysical method? If not, which one did they approximate?
7. What was indicated when three consecutive contacts were made in the shape of a triangle upon the skin?
8. What is the doctrine of "local signs"?
9. Was the doctrine of local signs confirmed or contradicted by Helson's and King's experiments? Explain.
10. What facts in neurophysiology tend to make the tau effect and related phenomena quite understandable?

CHAPTER 21

ADAPTATION TO TACTILE STIMULI

The sense of touch, although not by any means the most discriminating sense, is the agency by which we deal with a very important aspect of our surrounds. Touch was studied in the middle eighties by three men who independently discovered the nonuniformity of the skin surface to threshold or near-threshold stimulation. Blix, Goldscheider, and Donaldson discovered "sense spots" almost simultaneously. Max von Frey came a little later and soon became the leading authority in the field of touch sensation and remained so for a long time. Following the early work, relatively little experimentation has led to new ideas or understanding of this modality. The investigation that we are going to deal with in this chapter is an exception to this generality, however.

The Problem. Nafe and Wagoner, in 1941, concerned themselves with questions of (1) the relationship between weight and area of tactile stimuli and the amount and rate at which these sink into the skin; (2) times involved in adapting to these stimuli; and (3) rates of movement involved in contacting and removing stimuli from the skin required to produce pressure sensations. These questions imply that pressure sensations are dependent for their intensity and specific qualities upon weight and area of contacting objects, and that the observer senses not only contact but removal of contact of objects. These implications have their basis in everyday experiences. An individual certainly can sense the difference between the application of a sharp and an obtuse object to the skin. Likewise, the individual can tell when contact is made and when an object is removed from touching the skin. Furthermore, it is a relatively common observation to find that when something presses against the

skin for an extended length of time, the tactile impression of the object's pressure is diminished. This we may call adaptation. It was one of the initial concerns of Nafe and Wagoner to determine whether such a phenomenon should be classed as fatigue or as adaptation. For our immediate purposes, the factors bearing upon distinction need not be rehearsed. It is only necessary to say that the authors decided in favor of the phenomenon's being one of adaptation.

Apparatus and Stimuli. The apparatus Nafe and Wagoner designed lowered a weight onto the skin at a constant rate and recorded in a precise graphic form the sinking of an object into the skin under its own weight until it came to a halt. This apparatus consisted essentially of two parts, the stimulating system and the recording system.

The stimulating system was composed of (1) a member called a gondola with two forms of suspensions; (2) a stimulus weight holder; (3) stimulus disks; and (4) the precision control of lowering and raising the stimulus assemblage to and from the skin.

The recording system was composed of (1) a vertical slide with a stylus for writing on a kymograph drum, and (2) a pulley arrangement connecting the slide with the stimulating system.

The various parts of each system are illustrated first in Fig. 59, which is a diagram, and next in Fig. 60, which pictures the apparatus in action. Figures 61 and 62 are photographs showing the stimulus disk just touching the skin, and having sunk into the skin, respectively.

Let us first examine the stimulating system (shown in solid black). The gondola is shown as *A*. On its lower surface is a tube within which the stimulus disks *B* each fit and are held by a knurled setscrew. The gondola's upper portion is a vertical rod by which it is suspended. One suspension consists of a 6-g. counterbalance *C*, operating over a ¼-in. pulley *D*. The other suspension consists of a metal ring R 5½ in. in diameter from which three equally spaced elastic bands are attached to the gondola rod. This assemblage was adjusted in such a manner that the bands were horizontal and the gondola was held upright. The vertical tension exerted by the elastic bands, when the gondola was vertically displaced 10 mm. upward or downward,

was less than 0.1 g. at either extreme. During experimentation actual displacement was less than 5 mm. in either direction, therefore the tension factor was negligible.

The stimulus *weight* holder E was suspended so as to be lowerable onto the platform of the gondola, thus supplying the major

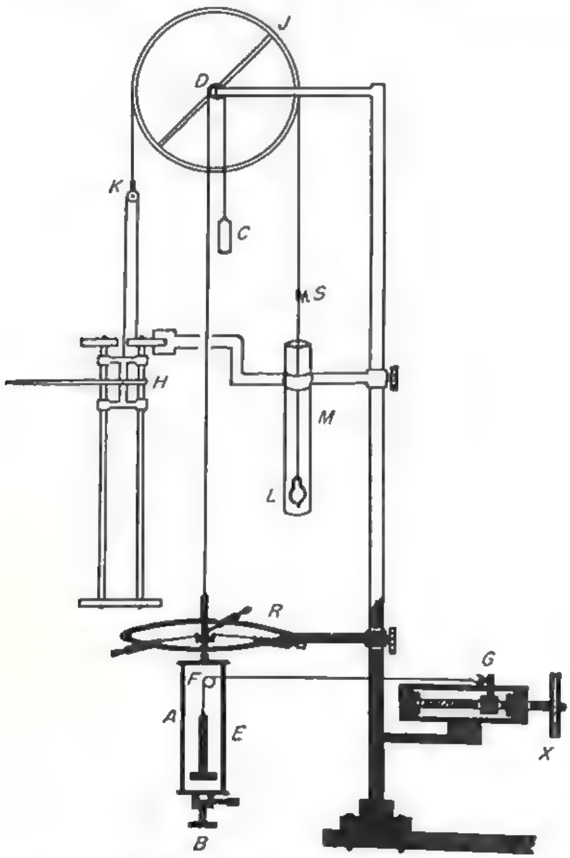


FIG. 59. Diagram of apparatus used by Nafe and Wagoner to investigate adaptation to pressure stimuli. It consists essentially of two major parts: (1) the solid black for stimulation, and (2) the outlined portion for recording. See text for detailed description. (*Nafe and Wagoner. J. gen. Psychol., Journal Press.*)

fraction of weight for the stimulus disk B. The suspension of E consisted of a silk thread run over a pulley F and attached to the device G for raising and lowering E. E and the weights it carried from time to time were lowered at the rate of 1.524 mm. per sec.

by a micrometer screw gear G, activated by a reversible constant-speed motor through pulley X and enclosed in a soundproof box. G was supplied with a switch automatically stopping the motor when the stimulus assemblage was delivered to the skin. When

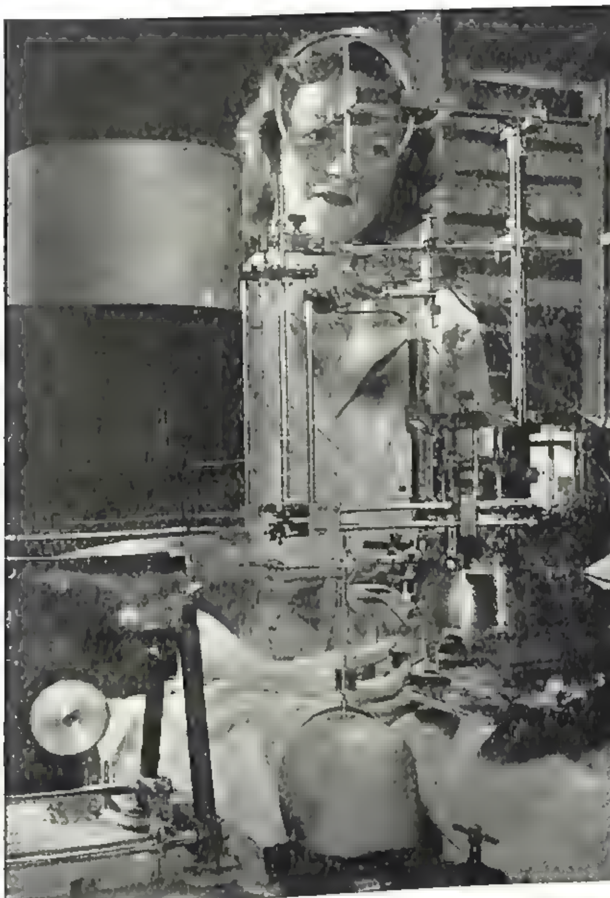


FIG. 60. The apparatus diagrammed in the preceding figure shown in action. (*Photograph by Paul Berg, St. Louis Post-Dispatch, Sunday PICTURES.*)

the motor was reversed, it lifted E and its weight. The gondola and the stimulus disk attached to it followed the E and weight upward to a position ready for the next trial. The stimuli used were 8.75, 17.5, 35, and 70 g. The weight holder E weighed 4.375 g. Used alone, it would provide a stimulus pressure of

4.375 g., so that another 4.385 weight was added to provide a stimulus of 8.75 g.

The recording system (shown in the unblackened part of the diagram) was related to the stimulating system in such a way that a vertical movement of the stimulus disk *E* and gondola *A* was magnified forty times. The slide *H* moved 40 mm. for every millimeter of stimulus-disk movement. This was accomplished

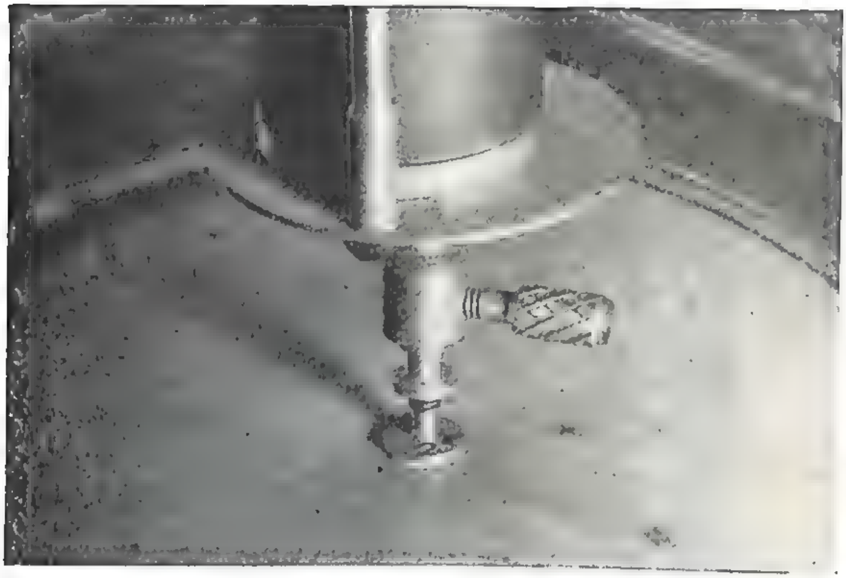


FIG. 61. Photograph of stimulus disk *B* of Fig. 59, shown just in contact with the subject's skin. (*Photograph by Paul Berg, St. Louis Post-Dispatch, Sunday PICTURES.*)

by having a 5-in. pulley wheel *J* mounted on the same axle with the small pulley *D*. The ratio of 5 to ¼ in. was 20 to 1. This ratio was multiplied by a factor of 2, by having a second pulley *K* attached to the slide-bearing end of the silk thread over pulley *J*. The opposite end of the thread carried a counterweight *L* immersed in a tube of damping fluid *M*. Pulley *K* was a tiny assemblage with jeweled bearings and it weighed 9.8 g. One end of the silk thread running over the pulley was attached to a stationary part of the slide holder; the other was attached to the slide *H*. Thus, as pulley *K* traveled up or down 1 mm., *H* traveled 2 mm.

Summary. The apparatus was so constructed as to lower and raise a stimulus disk of known effective weight onto and off the skin at a very slow and precisely controlled rate, and to register the vertical motions involved by a stylus writing on a precisely and slowly moving kymograph drum. The movement of the stylus was 40 mm. for every 1 mm. of stimulus disk. These results were achieved by having all moving portions of the

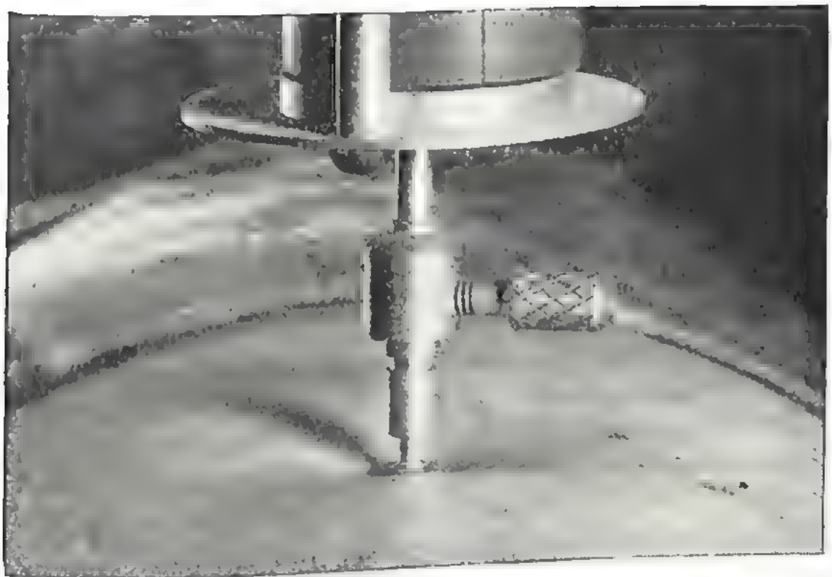


FIG. 62. Photograph of stimulus disk *B* of Fig. 59, shown after it had sunk into skin under its own weight. (*Photograph by Paul Berg, St. Louis Post-Dispatch, Sunday PICTURES.*)

recording and stimulating system very light in weight and carefully counterbalanced. Jeweled bearings helped to reduce friction. The stimulus disks were held by an elastically counterbalanced system (gondola, etc.) by which a stimulus disk was held at a fixed position above the skin until a displacing weight was lowered onto the gondola. It was this weight that provided the effective (uncounterbalanced) pressure of the stimulus disk on the skin.

Procedure. Two observers were used. One of these was naïve with reference to the aim of the investigation. The other was acquainted with the problem. During the experiment the

observer sat in a chair with the knee under the suspended stimulus disk. Various other areas had been tried and discarded in preliminary tests.

After application of the disk to the skin, the observer was told to report when adaptation was complete. That is, he was to report when the disk could no longer be felt. At no time during the tests could the observer see the weights and the area of the skin used. Various combinations of weights and areas of stimulus disk were used in random order, unknown to the observer. Trials were so spaced that no two trials were closer together than four times the adaptation period.

To measure the distance into the skin to which the stimulus disk sank, it was necessary that the measurement begin in each trial at the surface of the skin. Early tests showed that the precise level of the skin changes from moment to moment, and that any estimate of the point at which the disk barely touched the skin during a descent was far from accurate. Such error was obviated by having the disk in contact with the surface of the skin. With a very light pressure, the disk would rise and fall with the change in skin level, and this variation would also be registered on the kymograph record. Slight contact pressures of the disk with the skin were estimated to influence the depth readings as much as 0.03 mm. This amount was considered unimportant.

To effect the contact of the disk with the skin, a standard weight (0.285 g.) in the form of a tiny wire loop *S* was removed from the counterweight *L*, allowing the disk to descend slowly and rest upon the skin with a pressure of less than a gram. This pressure either was not felt at all, or else adaptation to it was complete in a few seconds.

Results. The relations between stimulus-disk area and weight and the indentation distance into the skin are indicated in Table 15. On the two subjects 1,370 records were obtained. This represented 30 measures for each subject for each possible weight-area combination, except two.

It will be noted that variations in area throughout the range of 1 to 32 (6.25 to 200.00) that was used made proportionately little difference in the distance of fall into the skin. For example,

with a weight (pressure) of 70.00 g., a shift from an area of 6.25 to 200.00 sq. mm. changed the fall from 3.67 to 2.11 mm. (1.74 to 1 decrease).

TABLE 15*

Area	Weight	Fall into skin
6.25	08.75	0.89
	17.50	1.59
	35.00	2.57
	70.00	3.67
12.50	08.75	0.83
	17.50	1.48
	35.00	2.45
	70.00	3.52
25.00	08.75	0.69
	17.50	1.28
	35.00	2.20
	70.00	3.34
50.00	08.75	0.52
	17.50	1.06
	35.00	1.91
	70.00	3.06
100.00	08.75	0.44
	17.50	0.85
	35.00	1.59
	70.00	2.65
200.00	08.75	0.33
	17.50	0.65
	35.00	1.18
	70.00	2.11

* After Nafe and Wagoner. *J. gen. Psychol.*, 1941, **25**, 325–361.

A shift in weight from 8.75 to 70 g. with area constant at 6.25 sq. mm. caused a shift in depth of fall from 0.89 to 3.67 mm. This means an increase from 1 to 8 in weight induced a shift of 1 to 4.1 in depth of fall (indentation). With a 200-sq. mm. disk the same increase in weight caused a shift in indentation of from 1 to 6.4 units.

The difference in varying area and in varying weight of the stimulus disk is shown graphically in Fig. 63. It will be seen

that as area is increased, the extent of fall is decreased somewhat; but that as weight is increased, the extent of fall is increased greatly.

The next consideration is the relation of fall to the touch experience. Naturally, as the disk sank into the skin, its rate of fall was at first more rapid than later, and the fall slowed up quite rapidly to a full stop. It was found that as the rate of fall became exceedingly slow, it reached velocity at which it became

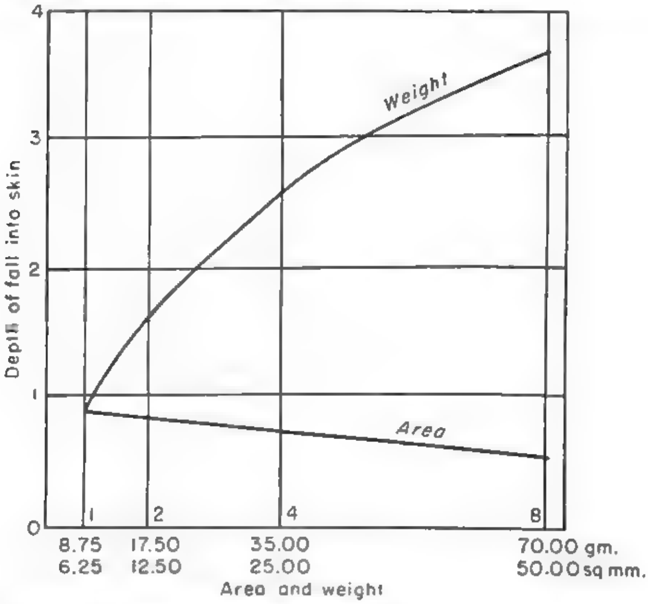


FIG. 63. Graph showing the depth that the stimulus disk sinks ("falls") into the skin, as either area or weight is doubled, trebled, quadrupled, and multiplied eight times.

ineffective in eliciting the touch experience. This we shall call the critical rate. This rate was not the same for all weight-area combinations of disk. It was found that as the *weight* of the stimulus disk is *doubled*, the rate of movement needed just to induce sensation (the critical rate) is reduced by one-third; or as the weight of the stimulus disk is *halved*, the critical rate is increased by a half.

One might have supposed that the critical rate of fall of the disk should have been the same in all cases. That this was not the case rests upon the fact that other factors beside the rate of

movement of the stimulus are involved in determining the nerve discharge giving rise to the touch experience. They are (1) volume of tissue moved, and (2) the extent to which it is moved. The "amount" of movement implies all three factors. No actual measurement of amount of tissue moved was included in the investigation, although the "falls" into the skin did measure the extent of movement. The authors believed that differences in the depth of fall of the disk into the skin are sufficient to account for the differences in the critical rates of movement obtained. This is believed on account of the following findings: (1) That there is a greater difference in extent of fall as the weight-size combination of the disks was varied from 8.75 g.–200 sq. cm. to 8.75 g.–6.25 sq. cm., than from 8.75 g.–200 sq. cm. to 17.5 g.–200 sq. cm., and (2) that there is a difference in the critical rate of the movement in the latter combination shift and none in the former.

QUESTIONS

1. What would you say is meant by adaptation?
2. What is adaptation usually attributed to?
3. What questions did Nafe and Wagoner concern themselves with?
4. Describe, in general, the type of instrumentation used by Nafe and Wagoner in their investigation.
5. What variables were used in their investigation?
6. Increase in what variable led to decrease in the depth of indentation (fall) of the stimulus disk?
7. How many subjects (observers) were used in the investigation? Should there have been many more? If so, why? If not, why not?
8. What range of weights and of areas of stimulus disk were used?
9. What was meant by critical rate of fall of disk into the skin?
10. To what did Nafe and Wagoner attribute tactile adaptation?

CHAPTER 22

STUDIES IN PROPRIOCEPTION

Proprioception is that form of sensitivity to body orientation and to application of gravitational and other accelerative forces by which the organism maintains posture, judges body motion, etc. Proprioception consists in two major components or divisions, the *kinesthetic* and the *vestibular*. The former division includes the receptors in muscles, tendons, and joints. These are activated by (1) forces applied to the body, and (2) positive muscular activity itself. The latter division includes the receptors in the nonauditory labyrinth of the ear. These labyrinthine or vestibular receptors are of two kinds, those in the otolith organs and those in the semicircular canals.

The function of the proprioceptive system is to enable appropriate posture, motion, and resistance to forces in the gravitation system, but it does not carry on these functions without the cooperation of *vision* and *touch.* Vision enables the organism to deal with motion in other bodies when the organism itself is stationary. Vision, of course, is also involved in the relative motions between the organism and other objects. Touch, with the experience of deep pressure, is a contributor to the complex of factors that function together in so unitary a manner as not easily to be analyzed.

Kinesthetic receptors were first discovered in the course of anatomical research. Selective extirpation procedures were used to determine the roles played by the several kinds of muscle, tendon, and joint receptors in the control of posture and body movement. The same was true of the sensitive structures of the inner ear. Anatomists have long known about most of the details of such structures.

Postural and movement reflexes were studied in connection with vestibular stimulation. Animal studies have predominated in this general line of inquiry. Clinical studies on humans have added their corroborations, but even yet few experiential investigations have been made on discriminatory response in this modality.

There are several natural forms of stimulation that activate vestibular receptors including (1) shift or maintenance of head position, (2) angular acceleration accomplished by rotation, and (3) linear acceleration produced in elevators, boats, planes, and in other devices. The vestibule is subject also to artificial stimulation, namely, by (1) heat and cold, (2) mechanical energy, and (3) electrical energy.

There are about five different types of experimentation that have been conducted on proprioception. They are the study of (1) receptor activity; (2) activity in pathways from receptor to cortex; (3) reflex responses in man and in animals; (4) direct discriminatory responses of humans and of animals; and (5) general experiential effects produced by stimulation conditions such as used in (3) and (4). These five types of inquiry are often referred to as experimentation at *different levels*. In much experimentation considerable attention has been and still is being given to the obtaining of stimulation *thresholds*. Such work stems from the belief that, once the absolute and relative sensitivities of the various sensory mechanisms involved are known, we shall be able to construct out of such material an understanding of proprioceptive behavior. Some doubt may be attached to this, for this form of inquiry by-passes the question of what the organism is trying to do on a given occasion and thus omits consideration of the various ways in which the organism utilizes sensory inputs in structuring behavior. Such inquiry, although fundamental, will not tell us why the organism can perform violent and rapidly changing movements of its own without developing "motion-sickness," whereas the same or even less violent motions externally applied to it may cause trouble.

Most motions imparted to the body by external means found in nature, such as in horseback riding, swinging, riding on trains, in automobiles, in wagons, and in boats, are very complex. They

may have (1) *up-and-down* components (as in elevators and boats), (2) *sidewise components* such as swaying or rolling (boats), lurching (trains), yawing (airplanes), and (3) *forward-and-backward* motion, as in the pitching of vehicles.

Various laboratory devices have been utilized to study the consequences of motion. Of the many specific applications there are a few essential ones such as (1) revolving chairs, (2) centrifuges, otherwise known as merry-go-rounds, (3) swings, and (4) elevators. Much of the earlier apparatus was quite crude, but present-day demands such as those involved in military aviation call for precise and intricate devices. One of these, for example, is a gigantic gimbal[1] arrangement in which an individual can be seated and moved in the three dimensions of space simultaneously and at controlled rates.

Griffith's Study of Dizziness. A number of years ago Griffith made a study of dizziness in which he revolved his subjects in a chair. The chair revolved at a uniform rate, except for the short time (half turn) involved in acceleration and the time for braking the chair to a stop. The task of the observers was to report upon the sensations that were produced in the experimental situation. The experimental period was fractionated into eight parts: (1) The period prior to the day's actual rotation; (2) the period beginning with the "ready" signal and ending with the initial experience of rotation; (3) the period of acceleration of rotation; (4) the period of constant rotation; (5) the period from the application of the brake to the beginning of afternystagmus;[2] (6) the period during afternystagmus; (7) the period following nystagmus and ending with next rotation; and (8) the period following the day's rotation.

For each of these fractional periods, the observers were required to give reports concerning the way they felt. It turned out that, although there were certain sensations common to sev-

[1] A gimbal is a mechanical mounting such as is produced by pivoting a ring within a ring. The first ring is also pivoted along an axis at right angles to the first. It is the principle of mounting used in a toy gyroscope.

[2] Nystagmus is the involuntary reflex jerking of the eyes. The motion is a horizontal one in which the eyes drift in one direction and quickly jerk back in the other. This is repeated at a more or less regular rate.

eral of the periods, the reports distinguished between the periods by indicating when certain kinds of kinesthetic sensations of tension, etc., originated. The reports also showed the relation between the presence of nystagmus and the experiences of movement after rotation had ceased.

The gist of the observers' reports indicated that many kinds of sensation result from rotation and constitute what is sometimes called vertigo, or dizziness. The sensations pertain mostly to kinesthesis from the muscles of the eyes, neck, limbs, and trunk. The experiences are those of strain, tension, pulling, aching, rigidity, nausea, etc.

The importance of the findings lies in the disclosure of the complexity and the importance of the bodily effects that result from rotation. It was found that the experience of dizziness and discomfort includes sensory components arising from what happens mechanically to muscles as well as to vestibular receptors.

Griffith's mode of attack was essentially different from the usual studies of proprioception. Customarily the phenomena studied, aside from the purely experiential outcome of vertigo or motion-sickness, are those of reflex character, such as ocular nystagmus or certain postural reflexes. Nystagmus has been studied both in man and in animals, and static postural reflexes mostly in animals. The implication underlying these studies seems to be that the motion effects are produced through stimulation of the vestibular receptors and the ocular reflexes. Although kinesthesis and tactile stimulation are credited with contributing to vertigo, so little work has been done in determining their roles that considerable doubt commonly exists as to their importance. Griffith's study is certainly a contribution that should not be ignored, for it suggests that much lies in the greatly ignored kinesthetic factor.[1]

[1] The study of the kinesthetic modality has been very scant from every standpoint. Study in this area is mainly represented by experiments on weight lifting. Such studies have contributed mainly to the development of psychophysical methods and to an understanding of human behavior with reference to implicit scales of perceptual value rather than to the complex experiential and other features of the muscle sense.

Travis' Study of Motor Response to Motion. A very different study regarding bodily motion deserves brief mention. It is Travis' study of response to a complex motion pattern supplied to a subject seated in a revolving chair. The motion of the chair is determined by a cam[1] driven at constant speed by an electric motor. Because of this the chair swivels back and forth in an irregular harmonic motion (see Fig. 64).

The movements of the chair made by the cam system can be compensated for in such a way as to maintain the chair in fixed position. This is done with a lever control system. When the chair does not move, a pointer on a target in front of the subject is still. The degree to which the subject is able to keep the chair still is the measure of the success in the task. An electric system that interrupts a circuit 8.2 times per second is active just so long as the chair is kept motionless to within 0.7 deg. At any time the lever control adjustments are insufficient to accomplish this, the interrupter is inactivated. It requires 73.5 sec. for the cam to make three revolutions, constituting one trial. If the subject were to keep the chair immobile to within 0.7 deg. the whole time, the record would contain 73.5 by 8.2 interrupter marks, or a score of 603. Thus the maximum score could be 603, and the minimum could be zero or thereabouts. The task was made just difficult enough by the 0.7-deg. tolerance for every subject to score and no subject to get a perfect score. The setup constituted a set of conditions calling for behavior somewhat simulating rudder-stick control in a plane.

While it may be thought that the task just described is primarily an eye-hand combination affair, it can be performed by blindfolded subjects. Such subjects, however, do not score quite so highly as when not blindfolded and allowed to watch the pointer. The performance can be taken as a way of measuring the ability of the subject to compensate for externally induced

[1] A cam is a mechanical device. It may be a thin plate or a wheel that revolves eccentrically on a pin or shaft. Thus if a member such as a pin or another wheel is pressed against it as it revolves, the member will have a reciprocal motion imparted to it. In the present case this reciprocal (back-and-forth) motion was applied to rotating a chair back and forth in a complex series of excursions.

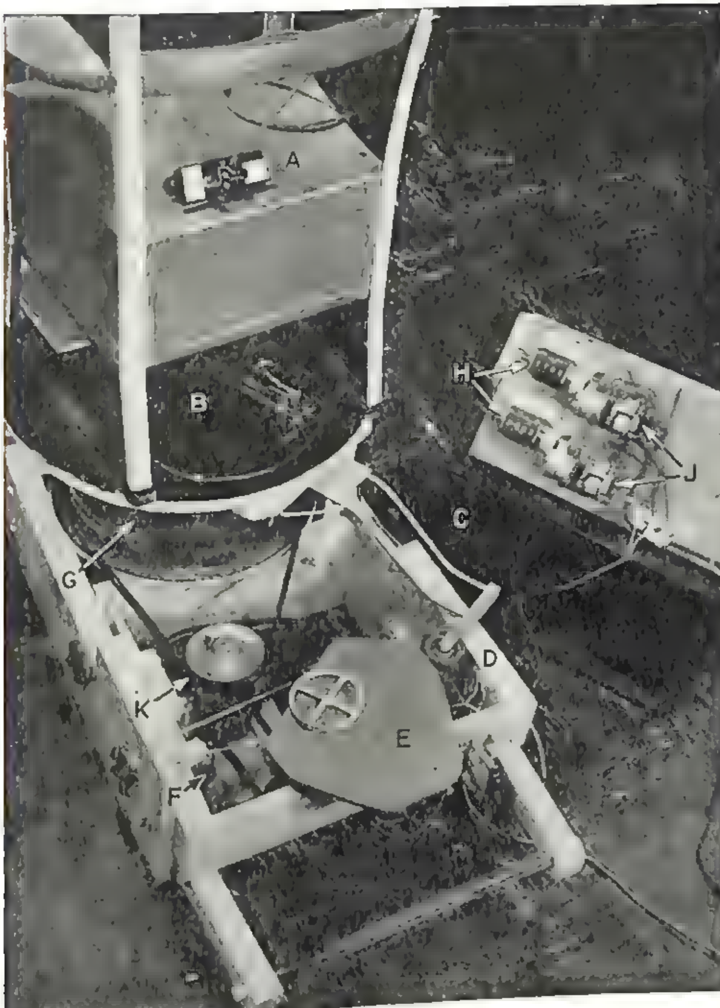


FIG. 64. Apparatus for studying response to rotation, consisting of a chair that oscillates irregularly from side to side if not compensated for by a "rudder stick" operated by the subject. A is response key used instead of a rudder for certain purposes; B, polarity switch; C, offset from chair to cam-follower; D, cam-follower; E, multiple harmonic cam; F, weight and pulley to maintain constant pressure on cam; G, axle; H, "correct" and "incorrect" counters used with A; J, relays for countermotors; K, speed reducer for driving mechanism. (*By courtesy of R. C. Travis.*)

movements. The task could therefore be utilized in a variety of proprioception experiments.

Travis used 100 college men and 99 college women as subjects in determining behavior in the experimental setup task just described. Each subject was given four preliminary trials to acquaint him with the conditions. The first was 15 sec. with eyes open; the next, 15 sec. with eyes closed; the third period,

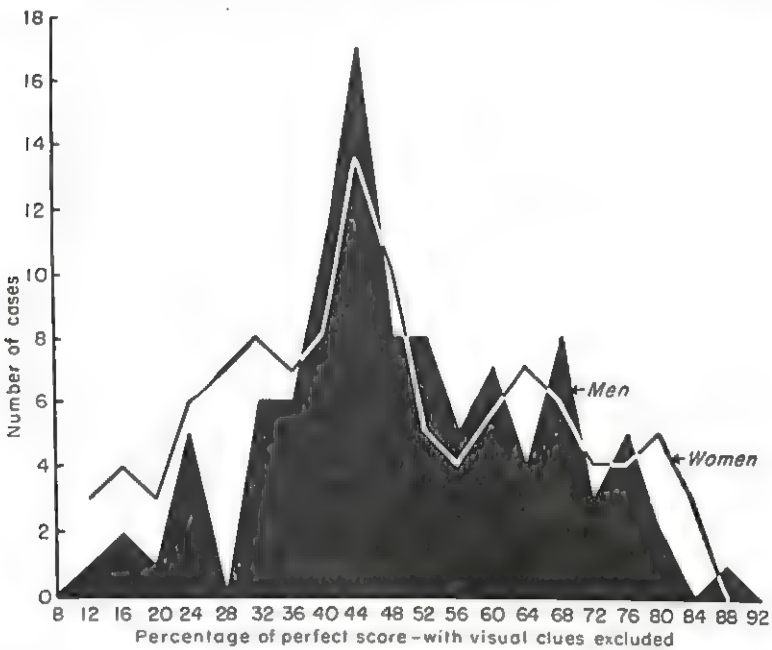


FIG. 65. Performance scores for men and women on Travis' rotating chair with subjects blindfolded.

15 sec. with eyes open, and the final period with eyes closed. Figure 65 shows the scores obtained by each group blindfolded. The results are given in the form of a graph in which the scores obtained are grouped in intervals along the abscissa, and the number of individuals obtaining the scores is indicated along the vertical axis. It will be noted that the most frequently obtained score by both men and women lies close to midway between zero percentage and maximum. Figure 66 indicates in the same manner the results when the subjects were not blindfolded. It will be seen that the value of the scores most frequently obtained

rose somewhat, moving the peak of the curve toward the upper range. Notwithstanding, the maximum possible score was not reached, nor was the highest score obtained much, if any, better than under blindfold conditions. The standard *deviation*, a measure of variability, is smaller in the open-eye trials than in the blindfold ones.

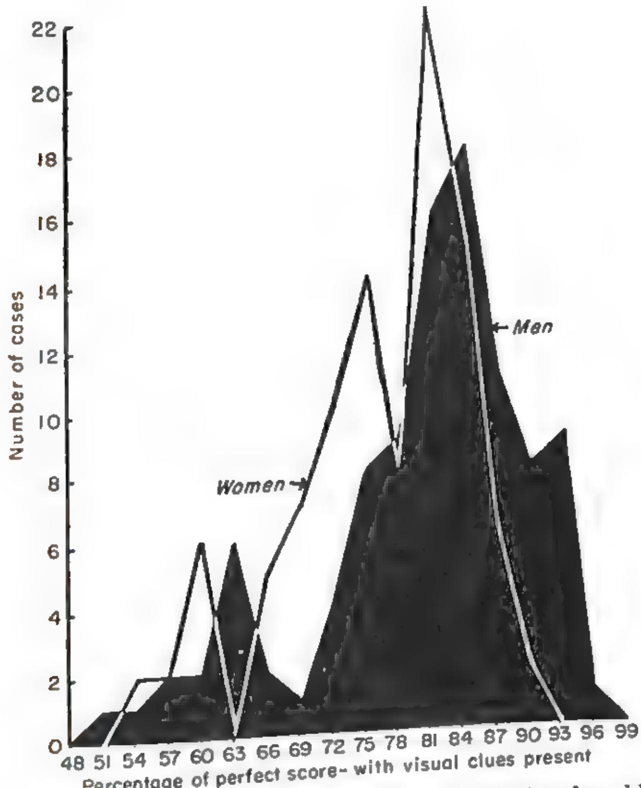


Fig. 66. Performance scores on Travis' rotating chair when blindfolds were not used.

One of the major outcomes of Travis' study was the evidence that he obtained for improvement in the task of bodily orientation to outside forces. Figure 67 indicates the rise in the scores for both men and women during a series of 10 trials, each of 73.5 sec. with 1 min. rest following each trial. The graphing of these results includes not only the curve showing the rise in percentage-perfection of scores but also the scatter of scores for both men and

women in each trial. So far, we have not presented any graphic representation of scatter in any study described. This graph of Travis', therefore, is significant as an example of this feature of data treatment. It will be noticed that there are only one or two cases in which definite reversal of trend is represented in the

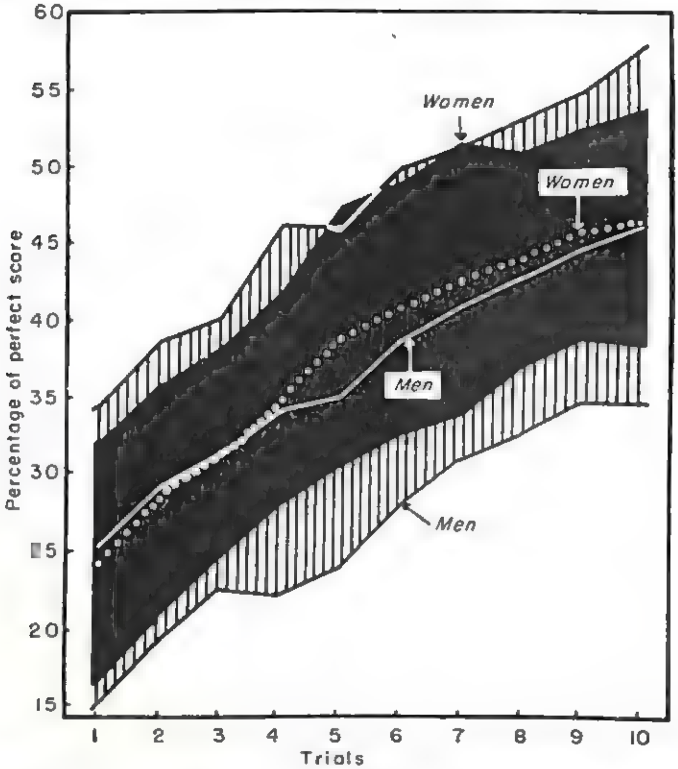


FIG. 67. Learning scores for 10 trials on Travis' rotating chair.

upper and lower limits of scatter. Some individuals have a higher score for the very first trial than some others have in the tenth trial.

The fact that one can readily learn to orient the body in relation to movements externally applied to it is a significant feature of the proprioceptive process. In order to make a more extended learning curve, the amount of tolerance in the apparatus might have to be less than the amount actually used so as to

increase the difficulty of the task. This means, of course, for some individuals the difficulty of the task might be made too great at the start.

Travis used the rotation chair also in a very different way. By discarding the "rudder stick," the chair was allowed to rotate just as the cam *E* and its mechanism determined. Since the cam was complex, it provided for a number of distinct rates of movement and one no-motion period of the chair. Eight different rates varying from zero to 2.3 deg. per sec. were produced. The task of the subject was to determine the direction of movement at all times. For this purpose two keys were provided, one to be pressed while the subject perceived the motion to be toward the left, the other while toward the right. In case of no perception of motion, no key was to be pressed. To record the subject's behavior were two counters *H*, each activated by a synchronous motor that was set in motion by its respective relay *J* when its key *A* was pressed. As the direction of the chair movement changed from right to left, or vice versa, a polarity switch *B* changed the circuit, so that one counter always recorded the correct responses, and the other one, the incorrect. The time required for a full rotation of the cam was 24.5 sec. Each segment of the cam acted, therefore, for a predetermined fraction of this period. These fractional periods varied from 2.0 to 4.8 sec. During the pressure of one of the keys, its respective counter remained in motion. The total possible score for either of the counters was 735, since the rate of revolution of the motors was 10 per sec. and the total time involved was 73.5 sec. (three revolutions of the cam). The actual score, however, was the number obtained by subtracting the incorrect counter score from the correct, and deducting from this the penalty score. The penalty score was given when the correct counter was activated while the chair was not actually in motion. The penalty score was given a value of 15, 30, or 45 points depending upon whether the perceived motion was registered during the first, third, first two-thirds, or during the full period of the zero motion.

The obtained score (raw) was divided by the perfect score (735) to obtain a percentage score.

QUESTIONS

1. What is proprioception?
2. What are its two major components?
3. What kinds of experimentation have been performed in proprioception?
4. What are some of the laboratory devices used to study the organismic consequences of motion?
5. What is nystagmus? What is a common way of producing it?
6. What was the objective of Griffith's study of dizziness?
7. Describe what Griffith did in his investigation.
8. What were his results?
9. Describe Davis' experiments.
10. What was the significance of his work?

CHAPTER 23

THE ORGANISMS' REACTION TO VERTICAL MOTION

It was pointed out in the preceding chapter that motion may be analyzed into several directional components, the vertical component or up-and-down movement being one of them. From a number of earlier studies it has become evident that the vertical component of any complex movement imparted to the body is very significant in producing motion-sickness in susceptible individuals. In light of this fact, a University of Rochester group[1] performed an extended series of experiments analyzing up-and-down motion into several of its aspects to determine their sickness-producing effects. To do this, a device resembling the common elevator was employed. The device was such that it could be raised and lowered alternately in an automatic fashion by setting the controls. The rates of ascent and descent, and also the up-and-down distances, could be varied.

Up-and-down motion on such a device can be thought of as a kind of wave motion. Hence we may call the elevator a wave machine, as did the experimenters themselves. The behavior of the wave machine may be pictured by means of wave forms in our diagram. Figure 68 contains three sets of wave diagrams that will be explained as we proceed.

The investigation represents a very good example of the comprehensive mode of undertaking research and is so extensive that only some of the major features can be described in this chapter.

The *first* study in the series of seven was designed to test the supposition that the time interval between the acceleration and

[1] S. J. Alexander *et al.* Wesleyan University Studies of motion sickness. *J. Psychol.*, 1945, **19–20**.

deceleration phases of a periodic up-and-down movement is a factor in inducing motion-sickness. To clarify the meaning of this statement, refer to Fig. 68 just mentioned. The waves in the first column represent three of the four motions used in the first study. The beginning of the first wave in the diagram represents an acceleration phase. The straight-line portion of the upward slope represents a period of constant velocity. In

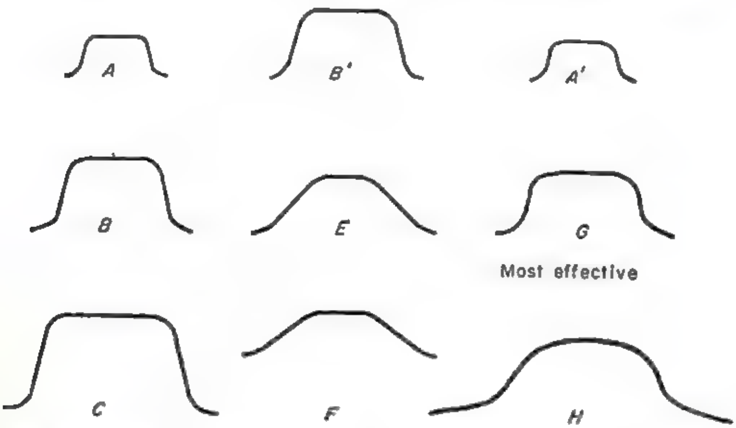


FIG. 68. Some of the waves used by Wendt and colleagues. *A*, *B*, *C*, constant vertical velocity of 400 ft./min.; *B′*, *E*, *F*, velocities of 400, 300, and 200 ft./min., respectively. For *A′*, *G*, *H*, acceleration varied, reaching a 400 ft./min. maximum at mid-point of wave. *A*, *B*, *C*, represent 32, 22, and 16 cycles per sec. with wave amplitudes of 4 ft., 7 ft., and 10 ft., respectively. Relative amounts of total work were 100, 69, and 50 per cent. *B′* is same as *B*. *B′*, *E*, *F*: 22, 22, and 20 cycles per sec. Amplitudes, 7, 5.83, and 4.66 ft. *A′* is same as *A*. *A′*, *G*, *H*: 32, 33, 16 cycles per sec. Amplitudes 4, 6.66 and 7.5 ft.

most cases this was 400 ft. per min. The tapering off of the slope to form the crest of the wave is the deceleration phase. The downward slope of the wave to the right of the crest is another acceleration phase. It is followed by a constant velocity period such as occurred during the upward travel of the elevator. Finally, as the slope tapers off toward the horizontal, we have another deceleration phase. This description applies, in principle, to every wave produced, and thus to every one in the figure. The periods of uniform motion represented by the straight-line upward and downward slopes, and the horizontal straight-line

portions of crests and troughs of the waves, represent periods between acceleration and deceleration. It will be noted that these straight-line portions of the waves increase from *A* to *D*. (Condition *D* is not represented in the diagrams.) To anticipate the results the experimenters obtained, it may be pointed out here that wave *C* produced the greatest amount of motion-sickness and wave *A* the least. This means that neither the longest nor the shortest intervals between acceleration and deceleration were the most effective in this respect.

In the *second* study, the duration of the period between accelerations was again varied, but the over-all duration of the cycle of up-and-down movement was kept roughly constant. That is, the over-all duration of the waves was about equal in the three cases, B', *E*, and *F*. The slopes of the waves that represented the uniform velocity reached at the peak of acceleration were varied. The peak velocity in B' was the same as the peak used in all the waves in the first study; namely, 400 ft. per min. Thus B' may be taken as a replica of *B*. In *E*, the velocity was 300 ft. per min., and in *F*, it was 200 ft. per min. In this study, B' produced the maximum motion-sickness.

In the *third* study, the acceleration rate was varied. The acceleration attained at the mid-point in the up-and-down motion was 400 ft. per min. in all four waves. But in A', *G*, *H*, and *J*, the rate at which this velocity was attained varied. In A', which was a wave form identical to *A*, the velocity was attained most rapidly. During the phase *m*, lying in the middle of the up-and-down movement, there was a constant 0.22-sec. period in which uniform velocity existed. It was the period of transition between acceleration and deceleration, or vice versa. In this study, A' produced very little motion-sickness, and *G*, representing less of an acceleration rate than A', produced definitely more motion-sickness. In fact, *G* in this study produced more than *C* in the first study, and, of course, *C* produced more than *B*, the very same motion as B' in the second study. Wave *G* possibly produced more than *H* in the third study, although the differences were not too clear.

Let us examine Fig. 68 to reach a clearer understanding of the variables employed in each of the three studies. In study 1,

pictured in column 1, the waves from A to C (D not shown) become greater in amplitude and in duration of crest and trough so as to involve greater amounts of time between accelerations and decelerations. In the three waves (B', E, F) in the second column, picturing study 2, the durations of the crests and troughs of the waves vary. Also the *durations* of the uniform up-and-down motions vary, so that by keeping the over-all duration of the waves approximately constant, the rates of motion are different. Hence the rates of ascent and descent constituted the significant variables. These rates were 400, 300, and 200 ft. per min., respectively. In the three waves (A', G, H), pictured in the third column, the rate attained when the elevator was at its mid-point m in ascent and descent was held constant at 400 ft. per min. The variable was the *rate* at which this mid-point velocity was attained during acceleration and slowed down from during deceleration. In A' the velocity was attained the most rapidly and held the longest. In the succeeding waves the acceleration and deceleration were slower and slower.

The *fourth* study was planned to determine how motion-sickness rates are related to waves whose top and bottom portions involve different rates of acceleration, as compared to a standard wave whose top and bottom portions are alike. All waves were of the same total energy, using a mid-wave velocity of 400 ft. per min. It was found that the symmetrical wave H' (identical to H in study 3) was more effective in producing motion-sickness than any of the asymmetrical waves employed.

The *fifth* study was designed to discover relation of motion-sickness rates to time of day. No consistent relation was found, although it was admitted that certain diurnal effects may exist.

The *sixth* study tested the incidence of motion-sickness in a 20-min. ride of blindfolded subjects on the wave machine for a possible relation to prior history of motion-sickness disclosed in a questionnaire given to each subject. In this study there was a positive relationship found between the effects in the tests and the susceptibility assigned from the questionnaire.

The *seventh* study was arranged to determine whether individuals becoming motion-sick manifest certain types of performance deficit. Tests were administered both before and after the

exposure to wave motion on the wave machine. Only unimportant effects between motion-sickness and ability to perform laboratory tests were disclosed. The duration of excitation for motion-sickness was brief in the experiments, since those subjects who became sick were taken off the wave machine as soon as this occurred. This brevity in the duration of the inciting conditions may have been a large factor in accounting for the lack of effects of motion-sickness upon subsequent performance. There may have been still other factors, so that no generalization seems possible from this particular study.

You have already been told about the variables in motion that were used in the first four studies and the other variables used in the final three. You should know something about the procedures used, particularly in connection with the wave machine.

Procedures. The subjects used in the studies were naval aviation cadets and/or naval officer candidate trainees. The three contacts that each subject had with the studies were: (1) the filling out of a personal history questionnaire; (2) the serving as a subject in one or another of the experiments, including the taking of certain performance tests; and (3) the reporting back regarding his history of airsickness during training, if a cadet.

As was said, the wave machine is something like an ordinary elevator in general appearance. The moving cab was a sealed, sound-shielded, temperature-controlled cab in a shaft permitting the production of up-and-down excursions of 18 ft., hence vertical waves of 18 ft. The well-illuminated interior of the cab was observed through large windows in its front as it changed direction at the level of the control booth. There was a two-way communication and recording system between the cab and the control booth. The cab was hydraulically driven, and controls for variation in rates of acceleration, etc., were provided.

The test period in the wave machine lasted for 20 min., or until the subject vomited. Subjects were not used for more than one test. Sufficient data were obtained for statistical reliability by using a large number of subjects and by virtue of the experimental design employed.

The experimenters used what they called a "three-variable system of counterbalancing" of the experimental factors involved.

Subjects of each of three grades of "susceptibility" to motion-sickness were used in each of the 10 test periods spaced an hour apart during the day. The grade of *susceptibility* of each subject to motion-sickness was assigned on the basis of the frequency and intensity of symptoms reported in the life-history questionnaire and on the basis of degree of experience in a list of motion situations. One-half or more of the subjects reported experience on at least 11 of the devices listed, and no consequent sickness. Such individuals were considered to be "nonsusceptible." The others had experience on less than the 11 devices and/or had been nauseated or had vomited. This group was divided into two equal parts on the basis of susceptibility. Such individuals were assigned to grades 1 and 2, and the nonsusceptibles to grade 3.

All possible combinations of stimulation and other variables were utilized. For example, at one test-period (let us say, 7:30 A.M.) all four wave forms used in study 1, partly pictured in column 1 in Fig. 68, were used. For each of the wave forms, one subject from each of the three susceptibility grades was used. Thus a total of 12 subjects were used at this one hour. The same arrangement was followed for each of the other nine hours, making 120 subjects per day. This tended to take care of the time-of-day factor as well as others. The experimenters tried also to counterbalance the various days of the week by assigning within each day 10 different combinations of "susceptibility" and wave form.

It was thought that the experimental design using counterbalancing possessed three special virtues: (1) It provided for more consistent results from experiment to experiment with relatively small numbers of subject than when the variables were allowed to occur as they might; (2) it provided for concurrent answers to the effects produced by each of the wave forms and other variables without doing separate experiments; and (3) the results were more valid for "real-life" situations than are the studies whose normal sources of variability are eliminated rather than being counterbalanced by statistical design.

As you may have gathered from the descriptions thus far, some subjects became sick and others did not. Naturally, there had to be a criterion for counting a subject sick. After the subjects

had ridden on the wave machine, they were assigned a number, 0, 1, or 2. Value 2 was given to those who vomited; value 1 was given to those who reported unmistakable nausea, and/or showed profuse sweating. The value 0 was assigned to all others. Most of the subjects in the 0 group had no untoward symptoms, but there were some of them who experienced dizziness, headache, slight nausea, pallor, or sweating which was less than profuse.

The ride on the machine was continued until vomiting occurred, or until 20 min. had elapsed, whichever came first.

The subjects were blindfolded during the ride. This seemed to quiet them and reduce the amount of shifting about during the ride. Clothing was reduced to the point where sweating would not ordinarily occur. The temperature was kept at 86°F. The rate of air circulation in the cab was enough to be perceptible on the face.

In summary, it may be said that these experiments formed one of the best of their kind for determining effects produced by well-controlled motion patterns imparted to the human subject. The wave machine itself is an arrangement that could be used for a number of other studies on human response to motion.

QUESTIONS

1. Into what components was vertical motion analyzed by the Rochester group?
2. What was the wave machine like?
3. What effect upon the organism was being studied in this investigation?
4. Who were the subjects used in the Rochester study?
5. How was susceptibility graded?
6. How were the reactions of the subjects graded?
7. How long did the subjects ride on the wave machine?
8. Was any relationship found between motion-sickness as produced in this investigation and the performance of certain set tasks?
9. What was the "three-variable system of counterbalancing"?
10. What was achieved by the experimental design used by the Rochester group?

CHAPTER 24

STIMULUS THRESHOLDS OF THE SEMICIRCULAR CANALS

One of the necessary things in the understanding of the factors underlying proprioception is the degree of sensitivity of the semicircular canals to movement. Apparently the form of motion to which the canals are sensitive is angular acceleration. Accordingly, the behavior of the canals has been tested by various devices imparting angular acceleration to the body. The conditions surrounding this type of study have varied, and with this has gone a range in the reported values for thresholds of sensitivity. These values have ranged from 0.2 to 4.0 deg. per sec. per sec.[1] angular acceleration. The lowest values reported are those from a study in which the subject was suspended in a tank

[1] At this point you will need to be clear regarding the distinctions between velocity and acceleration and the units of measurement for each. *Velocity* is the time rate of motion or displacement in a given direction. Thus to express velocity, you divide distance (or space) by the time involved in traversing it. Thus the formula will be $v = d/t$. Where v is velocity, d is distance, and t is time. For example, we say 60 mi. per hr., which is 60 mi./1 hr. Velocity may also be measured in angular terms (degrees, minutes, or seconds), when motion occurs around a fixed center. Then velocity is deg./time or, for example, 5 deg./1 min. or 5 deg. per min. *angular velocity.*

Acceleration, however, is the time rate of change in velocity, either in speed or direction. It is measured in feet, inches, or centimeters per second per second. One way of arriving at this *change in rate of displacement* of an object is to subtract initial velocity (v_o) from the final velocity (v_t) and divide the result by time [$a = (v_t - v_o)/t$]. For example, if you are moving at 15 min. per hr. and in 10 sec. you are traveling at 60 mi. per hr., you have accelerated 45 mi. per hr. in 10 sec. But, to express this in terms of what was done in 1 sec., customarily it is stated as "4.5 miles per hour per second per second," or 4.5 mi./hr./sec.[2]

of water with ears plugged and eyes blindfolded. The tank was rotated by a viscosity motor to ensure smooth acceleration. In most trials the subject was exposed to slowly *increasing* angular velocity. The rate at which he first reported apprehension of rotary motion was taken as the threshold.

Throughout experimentation in this area, investigators have been dependent upon a very limited set of indicators by which to ascertain the sensitivity of the semicircular canals. One of these, as was pointed out in an earlier chapter, has been ocular nystagmus; the other has been the experience of motion. Measurements of nystagmus have varied greatly on account of the accuracy and delicacy of recording eye-movements.

Depending upon this factor, nystagmus has been variously observed. It has sometimes been recorded as a phenomenon that subsides and disappears not long after the cessation of the body motions that induce it. At other times, it has been recorded as a response persisting for long periods (minutes) after the termination of the external conditions inducing it.

There is no known necessary concomitance between sensory phenomena and the absence and presence of nystagmus. To think of motion experiences in the absence of nystagmus is plausible. On the other hand, if nystagmus is as persistent an affair as just suggested, incipient forms of nystagmus might underlie threshold experiences of movement without being recognized as their basis. To say the least, our knowledge of sensitivity to movement might be enhanced by having other indicators in addition to outright kinesthetic experience and nystagmus.

A phenomenon that has proved useful for this purpose is what has come to be called the "oculogyral illusion." It is the apparent motion of a visually perceived object following angular acceleration of the body. It is presumbly the result of the angular acceleration on the semicircular canals. One of the very best ways to demonstrate the nature of this phenomenon is to rotate a subject in a darkroom. There should be a light rotated with him and stopped when he is stopped. When this is done, the light does not appear to the subject to stop and remain still with him. When he stops, the light seems to move in the direc-

tion opposite to which he and it were previously rotating together. Later this apparent motion may slow down and be substituted by a movement in the reverse direction. Since it takes only slight acceleration to bring about the appearance of the oculogyral effect, it was thought to be a likely indicator for measuring thresholds of sensitivity to angular acceleration.

The Problem. A group[1] at the Naval Air Training Station, Pensacola, Florida, undertook to make new determinations of the threshold for angular acceleration using the oculogyral effect just described and the physical facilities at their command. The problem was to supply subjects with very low rates of acceleration and deceleration and to record the appearance of the oculogyral effect in connection with a lighted target that moved with them.[2]

The Apparatus. This consisted of a human centrifuge (a merry-go-round) constructed primarily for high velocities, but also quite appropriate for the low rates to be employed in the present study. The centrifuge was of the inertia-wheel type. That is, the uniformity and smoothness of motion was supplied by the inertia of a huge flywheel. In this case, the flywheel weighed 22 tons and was propelled by a friction drive from a gasoline engine. The platform of the centrifuge, on which the subject rode, although coaxial with the flywheel, was independently suspended. The platform could move when the flywheel was still, or could remain still when the flywheel was in motion. To be set into motion by the flywheel, it had to be engaged by a clutch. The combined platform and flywheel could be disengaged from the engine by means of a second clutch. The platform occupied one story in a building and the flywheel the story below it. The operator of the platform and flywheel was stationed in a separate room.

For the present study, after the flywheel was in motion and the subject's platform clutched to it, acceleration was obtained by increasing engine speed. Deceleration was provided by

[1] A. Graybiel, *et al.* Stimulus threshold of the semicircular canals as a function of angular acceleration. *Amer. J. Psychol*, **61.**

[2] Keep in mind that to have the oculogyral effect appear, one does not have to stop, one has only to change rate of motion.

disengaging the engine from the flywheel by the second clutch. Friction of the system very slowly brought the system to rest. This, at times, required several minutes depending upon the rate just before deceleration was begun. The subject was fastened into a bucket seat facing the axis of rotation with his head 2 ft. from the axis. The subject was rotated only to his left. His head was inclined slightly forward (20 to 25 deg.) and fixed by a biting-board.[1] In front of him was a lighted star pattern of six radial lines. There was a fixed positional relation between the star and the subject's head, so that any movement of the star that the subject might see would be apparent, or the so-called "illusory" type.

The velocity of the centrifuge was indicated by a tachometer dial. A written record of this was made by a writing lever forming a continuous line on the record. The elevation of this line represented the velocity. A signal lever also made a mark on the record at every revolution of the platform. Another writing lever of the recorder was activated by a push button when the subject wanted to indicate apparent movement of the star (the oculogyral effect).

The subject was housed in a plywood box or room so as to eliminate air motion as a cue. This, of course, eliminated visual cues that might otherwise exist. The subject wore earphones. These accomplished two purposes. They were used as indicators of 20-sec. intervals, the use of which will be disclosed later. Because of the steady tone introduced into the earphones, extraneous noise and auditory cues were brought to a minimum.

The Forces Acting. The physical features involved in the situation include radial acceleration, angular velocity, and angular acceleration. Actually, the only desired factor was angular acceleration, but the other two factors were found not to be involved in ways that would disturb the subject's observation and the experimenter's analysis of results. *Radial acceleration* is the constant change of direction of movement of an object at any distance from the center of rotation of a disk, wheel,

[1] A biting-board is a rigid stand carrying a plate or wooden stick upon which the subject bit his teeth and enabled him to maintain uniform head position and place his head in the same fixed position upon repeated occasions.

centrifuge, etc. This acceleration becomes effective as centrifugal force. Since the subject's head was about 2 ft. from the center of the centrifuge-platform's axis, some radial acceleration was involved. At the very maximum rate of rotation, 18 r.p.m., the outward force, due to this acceleration, was 0.22 G (*i.e.*, 0.22 of the effect of gravitational force). This force acted horizontally, and it and gravity together exerted a pull on the individual so as to make him feel tilted backward as he faced the center of rotation of the platform. The subject, of course, was fixed in a rigid, almost upright position by his grip on the biting-board, or he would have leaned inward toward the center of rotation and felt no outward pull such as just described.

A striking feature that accompanied the subject's feeling of leaning outward was that the star originally seen at eye level tended to be located *above* the horizontal. This effect occurred only at the higher rates of rotation, *i.e.*, above 10 r.p.m. Acceleration and deceleration increase and decrease the effect just described, so that the subject tended to see the star rising above or descending back to a level horizontal with the eyes. However, with the low rates of acceleration and deceleration used, any such subjective vertical movement of the target was very slight and formed no dependable cue for the angular acceleration under study.

The second factor involved in the conditions of the experiments, *angular velocity*, was counted as of no significance.

Procedure. Each trial began with the subject with closed eyes and prepared to use his push-button signal. The movement of the platform began abruptly at 1 or 2 r.p.m. The flywheel was started and obtained this rate before the platform was clutched in. From the point of clutching in, acceleration began and took the platform slowly up to a rate of 5 or 10 r.p.m. Naturally, it was not the actual number of revolutions that was dealt with in this study, but rather the *changes in rate of motion*. More about this factor will be described presently. After acceleration was carried to the point of maximum desired velocity, a constant speed was maintained for at least 3 min., so as to allow for recovery from accelerational aftereffects. After the elapse of about 3 min. or more, the subject began to receive in his ear-

phones a signal every 20 sec. This was in the form of a sudden and transient change in the masking noise already mentioned. In response to each signal, the subject opened his eyes and was given 5 sec. to report any apparent motion of the star, if any. This report was made by pressing the push button, once for motion to the right and twice for motion toward the left. When no motion was perceived, no signal was given.

Each run consisted of a series of three or more trials in each of two general periods—one in which the velocity of the platform rotation was increased and one in which it was decreased.

Results. The data consisted of 1,580 twenty-second response periods with the computed average positive, negative, or no-acceleration values for each. These data were used in two ways, only one of which will be described here. The utilization of the data consisted in plotting the positive and negative acceleration values against reports of left-movement, right-movement, and no-movement of the star pattern.

The subjects' reports tended to indicate apparent movement in the same direction as the actual acceleration itself. Nevertheless, the relation between the reports and the existence and amount of acceleration or deceleration was quite irregular. This suggested that the 20-sec. periods might be too short as intervals for judging. As a consequence, four 20-sec. periods were combined to give a longer time unit. The *method of moving averages* was also resorted to. This consisted in obtaining consecutive 80-sec. units by dropping off the first 20-sec. period of the group of four such periods just used and adding on the next 20-sec. period in the record to form the next 80-sec. unit.

The records consisted of a continuous-line tracing whose elevation from base line represented velocity of the centrifuge. The rise and fall of the tracing from base line, therefore, indicated acceleration and deceleration (plus and minus acceleration), respectively. The height h of the tracing at each 20-sec. period was measured. The differences d between the heights in each period indicated whether acceleration had occurred or not. If all four d's in an 80-sec. period were equal, the acceleration was constant. If they all differed in one direction, the acceleration might be rising (+) or falling (−), depending upon the direction

of the differences. If some differences were in one direction and some in another, the average might come out positive or negative or zero.

The results of plotting this information are shown in a series of three graphs (Figs. 69, 70, and 71). In each, the abscissa represents the mean change in angular acceleration, positive and negative from zero. The ordinates represent the percentage of

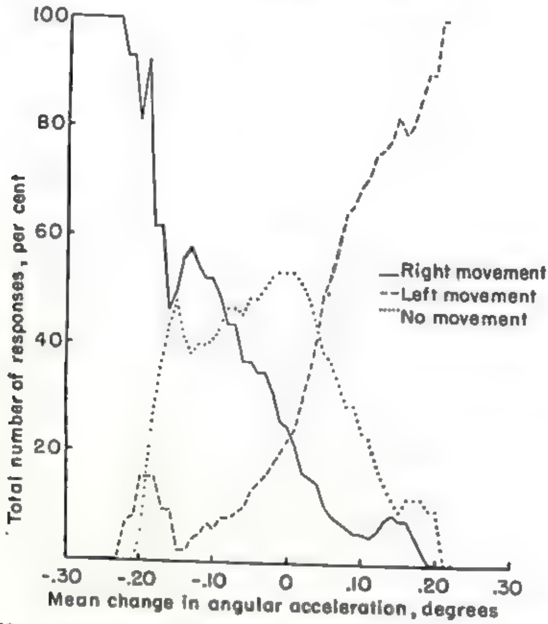


FIG. 69. Judgments of motion based on the oculogyral effect.

the total number of responses. In the first figure the right-, left-, and no-movement responses are indicated by curves shown in solid, broken, and dotted lines, respectively. The curves in the second and third figures can be understood from those in the first. It will be noted that most of the responses indicating movement to the right were made when acceleration was negative; most of those indicating left movement, when the acceleration was positive. Likewise, most of the no-movement responses occurred when no acceleration existed.

Figure 70 represents a simple kind of smoothing technique. A Gaussian curve (the normal probability curve), whose height was determined by the height of the empirical curves and whose

spread (three standard deviations) was determined by the spread of the data on the abscissa, was used. When this Gaussian curve was placed on the graph for both positive and negative accelerations, the fit was as good as could be expected by any ideal curve. The two curves crossed at about 23 per cent instead of the expected 33.3 per cent. In Fig. 71, the same curve-fitting technique was applied to the data representing no-movement

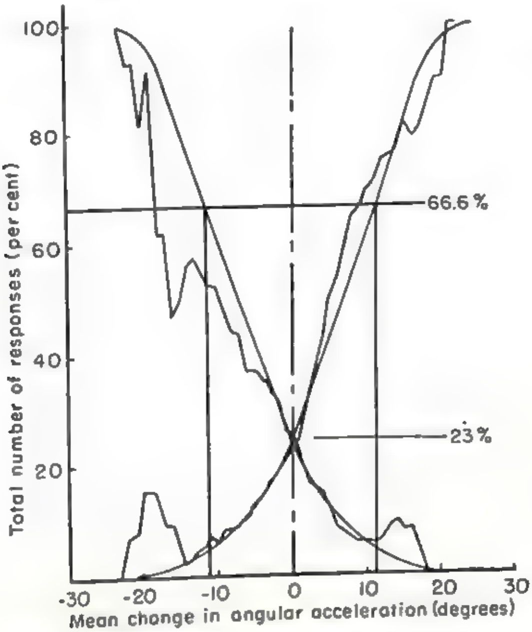


FIG. 70. Gaussian curves applied to the raw performance curves for right- and left-movement responses of Fig. 69.

responses, yielding curve *C*. The height of this curve was set so as to include 54 per cent of the no-movement responses when no acceleration had occurred. This was done so that the two other curves, each representing 23 per cent ($23 \times 2 = 46$), plus the height of the present curve would equal 100 per cent. A good or better fit was obtained by curve *B*.

Since there were three choices, the level of probability to use for threshold is 66.6 per cent; *i.e.*, halfway between chance (33.3 per cent) and full accuracy. If one were to note the 66.6 per cent level of probability on the raw curves in Fig. 69, the values of positive and negative angular acceleration would not

be the same. They would be about 0.9 degrees per second per second (0.09°/sec.^{2}) for positive angular acceleration and about 0.18°/sec.^{2} for negative angular acceleration. The values on the smoothed curves, of course, are the same for both positive and negative values, namely, 0.12°/sec^2.

It was reported as somewhat surprising that such low values were obtained for the thresholds of acceleration. The values

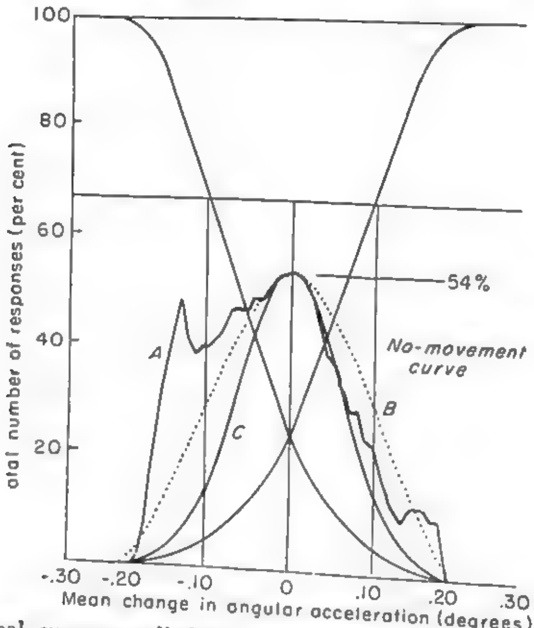


FIG. 71. Ideal curves applied to the raw performance data for the no-movement responses shown in Fig. 69.

are lower than any reported elsewhere. The combination of experimental conditions, psychophysical procedures, and the nature of the indicator might rightfully be expected to reveal greater sensitivity of the organism than disclosed by a number of the other investigations reported in the literature.

QUESTIONS

1. What is acceleration?
2. Distinguish between linear and angular acceleration.
3. What is the "oculogyral illusion"?
4. Describe the Pensacola apparatus used in the present experiment.

5. What is radial acceleration, and what is its mechanical result?

6. What was the stimulus for inducing a feeling of body tilt in the present experiment?

7. What forces acting in the experiment were so small as to be ignored?

8. What is the "method of moving averages"?

9. Describe how the "oculogyral illusion" was employed as an indicator of the individual's sensitivity to motion.

10. What was the threshold value obtained in this investigation in relation to the thresholds found in other investigations?

PART V

PHYSIOLOGICAL PSYCHOLOGY

CHAPTER 25

PARALLELS BETWEEN REFLEXES AND SENSATIONS

Two of the organismic consequences of stimulation are motor activity and sensory experience. One of the most immediate motor consequences is the type of reaction that is generally termed a reflex. The immediate experiential responses are perception. It is quite a common tendency in thinking about these two forms of response to view them as different in every respect. This view is not necessarily the fruitful one to have. There are certain phenomena that remind us that whether it is movement or sensory experience that results from some kind of physical energy impinging on sense organs, the immediate channel for this effect is the pathway from the sense organ to the central nervous system. The dividing path for the two kinds of reaction may be in the spinal cord or it may be high up in the central nervous system, depending upon which sense-modality is involved. Some sensory pathways lead to the central nervous system over *spinal nerves* and others over the *cranial nerves*. Touch, for example, is mediated over spinal nerves, and vision over cranial nerves (the optic nerves).

Since there is this common pathway to start off with, there is the possibility that both movement (motor reactions) and experience (perception) will have certain common characteristics. If this should be found to be true, our thinking about the nature and organization of behavior would be quite different than if little or no similarity in the patterning of movement and experience were found to be the case.

It is the aim of this chapter to describe some phenomena that will demonstrate the characteristics of certain reflexes and sensations that run parallel to each other. The reflex that we shall de-

scribe is one of the *light reflexes*[1] of the pupil of the eye. In connection with these reflexes, the human response to light intensity that is called brightness perception or *brightness discrimination* will be described. The study involving this comparison was made a few years ago by Bartley.

Perhaps we had better start with the description of sensory effects—what happens in brightness discrimination. Many years ago, Fechner performed a little experiment that disclosed unexpected results which ever since then have been called Fechner's paradox. He arranged a setup so that the two eyes could be stimulated independently. What he did was simply to provide a disk of light that could be seen by one eye and not the other. He also provided a second disk of similar size for the second eye that could not be seen by the first. To do this, all that was necessary was to put these two disks of light side by side with a little separation between them and then put a shield or septum out from the eyes, as is shown in Fig. 72. This cuts off the light from the disks in the way just described.

When your two eyes look at a single object, they each point toward it. We say they converge. But, in Fechner's set up, the two eyes did *not* converge. Each eye pointed toward its own disk. This resulted in the two disks acting as if there were only one, so far as perception was concerned.

The observer sees only a single disk just so long as the eyes point as just stated. With such an arrangement, it is possible to vary the intensity of the light in each eye independently from

[1] There are a number of pupillary reflexes. Some of them are responses to light, one of them is a response to darkness, and still others are responses to other types or aspects of stimulation. It has been discovered that response to a brief, dark interval injected into otherwise continuous illumination evokes pupillary reflex when the brain is so damaged as not to respond to light pulses. The pupillary near-reflex is the constriction brought about by viewing objects close-up. Pupillary apertures also reflect the effects of drugs, fright, etc.

Pupillary behavior is regulated by two sets of muscle fibers in the iris, one set for closing or constricting the pupil, the other for opening or dilating it. These two sets of fibers are activated by the two parts of the autonomic nervous sytem, the sympathetic and the parasympathetic. Hence the nervous pathways for the two sets of fibers are quite distinct throughout most of their courses.

zero to some desired level. Fechner began by presenting a bright disk to one eye and no disk to the other eye. He then wondered what would be the resulting brightness if he were to add to the original stimulation by presenting a dim disk to the second eye. Since the observer sees a single disk, whatever is done to either stimulus disk ought to affect the brightness of the disk seen by the observer. Naturally, one would expect that the observer would see a brighter disk because of the addition of the light via the second eye. This is *not* the result that occurred.

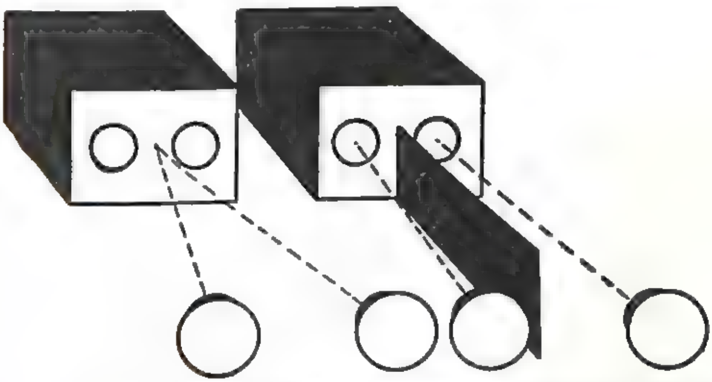


FIG. 72. Schematic representations of arrangement (on left) for both eyes viewing one or two disks of light. When only one disk is viewed, the broken lines from the eyes should converge on it. The right-hand arrangement indicates conditions when two disks are seen as one, since each eye is pointing to a separate disk. In this arrangement the intensity of the disks can be altered independently.

The brightness lay somewhere between those resulting when the two eyes were stimulated separately by their receptive disks. No wonder this was called a paradox!

Fechner did not make an extended quantitative study of this paradox. In so doing, he failed to disclose the further important features of this general situation. Had he varied the intensity of the second disk through a range up to the intensity of the disk used for the first eye, he would have found that when the intensities of the two became more nearly alike, the dimming effect of using the second disk diminishes and disappears. Finally, when the two intensities are equal, a very definite summation effect results. Thus we can say that when the stimulation of each

eye is equal, the combined effect is much greater than when only one eye is stimulated. Keep these sensory effects in mind as the reflex effects of using the same stimulus setup are described.

Apparatus. The apparatus for determining pupil size was a 16-mm. motion-picture camera supplied with infrared film. The room was dark except for the two stimulus disks and the source of infrared radiation. This was provided by two metal lamp houses about 10 in. from the observer's face with their axes about 30 deg. from the line of regard, as is shown in Fig. 73. In each

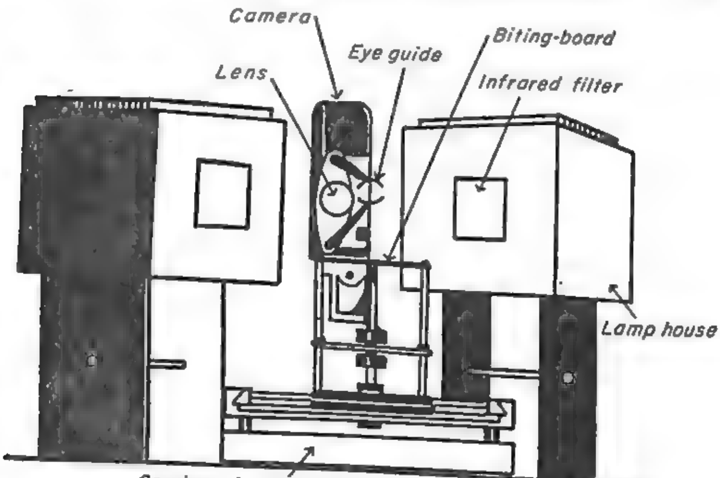


FIG. 73. The experimental setup for infrared photography of the pupillary reflex. The stimulus source, such as shown in Fig. 72 (right), is placed above camera.

lamp house was a powerful photoflood lamp. The exit for the radiation was through a 3 by 3-in. Corning-glass infrared filter and another filter that eliminated certain other undesirable wave-lengths.

The motion-picture camera was run at only a few frames per second, since only a few pictures under each condition of stimulation were needed.

Procedure. The subject was dark-adapted for a standard length of time, then one eye was exposed to its disk at a low level of intensity. Photographs were taken during this exposure. Then another somewhat higher intensity of the disk was used, and more pictures of the pupil taken. This was repeated for a

number of progressively increasing intensity levels of disk until a desired maximum intensity was reached. This maximum intensity was maintained for the first disk, while the second eye was exposed to its disk. The series of intensities used for the second disk were the same as those for the first. Pupil pictures were taken for this series as for the first.

When the procedure was complete, and the photographs developed, they were projected to a standard distance and the diameters of the pupils were measured frame by frame.

Results. With the information (pupil diameters) at hand, the experimenter was ready to plot the relation between pupil diameter and stimulus conditions with one eye and with two eyes stimulated. Before we examine the results, a fact or two regarding pupillary light reflexes should be mentioned. If one eye is exposed to light and the other is kept in darkness, not only will the pupil of the illuminated eye contract but also the pupil of the other eye. The two pupils contract about equally. This is called the *consensual reflex* in the second eye. Thus the consensual light reflex was involved in our procedures.

Figure 74 shows the results of illuminating one and then both eyes. The graph contains two curves. One of these represents the pupillary results when the light used was all cast on one eye. The other curve represents the results when first one and then both eyes received this light in the manner already described.

The curve for the one-eye then two-eye stimulation indicates that, with increasing intensity, the pupil diameters of the eyes became smaller and smaller just so long as the light was presented only to one eye. The extra perpendicular line in the graph indicates the point at which light was introduced into the second eye. From this point on, the first major effect is that of dilatation (increase in diameter) of the pupils. This is as if *less* light rather than *more* light were being presented to the eyes. Soon, however, this paradoxical effect reaches a maximum and declines. Finally, the combined effect of exposing both eyes to light results in greater pupillary constriction than previously reached in illuminating one eye only. Actually the ultimate effect is greater than when the same total amount of light is supplied to the one eye only. You will readily see that the same pattern of quantitative

effects are produced in the pupillary reflex as in perception. Both are paradoxical in the same way, *i.e.*, under some conditions *more* light produces *less* effect. Pupillary behavior parallels Fechner's paradox.

Three interpretations might be placed upon the results just described. It might be taken that (1) the pupillary reflex is simply somewhat more complex than generally supposed; (2) the pupillary reflex is under elaborate control of the higher centers,

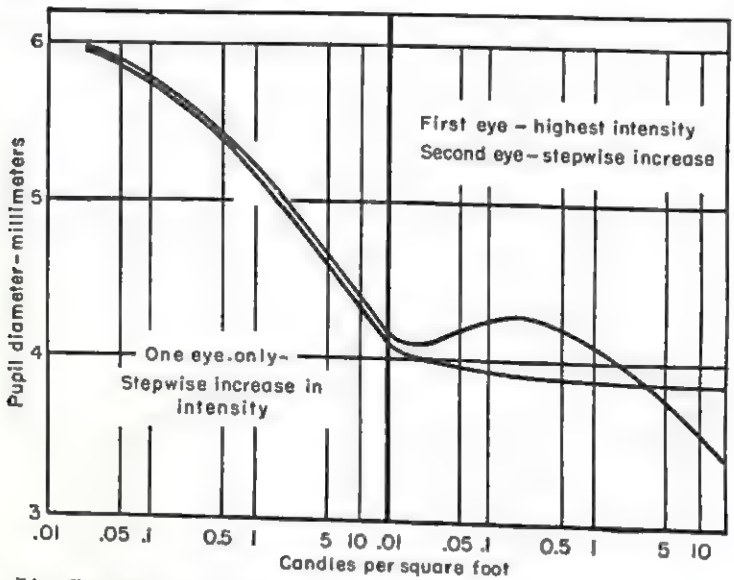


FIG. 74. Pupillary response, first when only one eye is illuminated, then when both are illuminated. The initial illumination for each eye is weak and is increased a step at a time.

as many other reflexes are taken to be, and thus reflects this control by paralleling the behavior of higher centers; and (3) reflexes and sensory discriminations are not essentially different in plan, but are both the expression of the behavior of networks of neurons. Whether consciousness results from the activity of one set of neurons and not from another might be quite incidental. It may have little or nothing to do with the patterns of activity involved.

To say the least, the mode of inquiry represented in the present study exemplifies a possible means of stimulating fruitful thought regarding ways of viewing organismic behavior.

QUESTIONS

1. What is meant by "parallels" between reflexes and sensations?
2. What does the pupil of the eye do in response to light?
3. What is Fechner's paradox?
4. What is a consensual pupillary light reflex?
5. Under what conditions does an observer see a single object when two are presented?
6. Describe the instrumental conditions used for studying the pupillary light reflex.
7. Describe the stimulus conditions for making this study.
8. If light in the stimulus to a single eye were made twice as bright, would the pupil constrict to half diameter, constrict less than this, or would it dilate?
9. Was a parallel between reflex and sensory behavior demonstrated?
10. What interpretations could be given to the results?

CHAPTER 26

THE CORRELATION BETWEEN NEURAL AND PERCEPTUAL EVENTS

One of the ways by which we can seek to understand the behavior of the organism is by attempting to discover certain relations between perceptual and neural events. Actually, overall behavior is a form of interplay *between* the whole organism and its surround; but to understand this, we have to inspect the mechanisms whereby the organism behaves as it does. These mechanisms or means-to-ends lie *within* the organism. The types in which we are interested in this chapter are called physiological mechanisms, since they are expressed in the activity of body tissue.

Since an understanding of physiological mechanisms is a fundamental objective in psychology, the description of an investigation disclosing the nature of one such mechanism is in order. The investigation we are choosing disclosed that a certain visual phenomenon was associated with specific features of the optic-nerve discharge.[1]

A few general considerations regarding response to light must be reviewed before we proceed with the description of the investigation in question. One of the supposedly most simple visual stimulus conditions is the presentation of a "flash" of light to the eye. Let us call this physical event a *photic pulse*, or a light pulse, and the experiential response, a *flash*.

A single, short light pulse does not always give rise to an equally short and simple experience. A pulse of a few millionths of a

[1] The optic-nerve discharge is the total group of nerve impulses that is set up in the neuroretina and sent into the optic nerve to be conducted to the brain when the eye is stimulated with light. The flow of impulses up the optic nerve is far from uniform, and their temporal distribution is a factor of considerable significance in understanding neuroretinal processes.

second, when experienced at all, is seen as a flash of much greater duration. In fact, no occurrence is experienced as being so short as a few millionths of a second.

Not only do experiences and stimuli differ in duration, but experiences may not possess the same quantitative *pattern* as the physical events that elicit them. Whereas the onset of the photic pulse may be abrupt, the experiential consequence may be noticeably tapered. The termination of the pulse may be as abrupt as its onset. In contrast, the sensation may possess a noticeable taper toward termination.

Sensory physiology and psychology have made numerous studies on the characteristics of the stimulus and its associated sensory response. Under some conditions the sensation is a very complex affair consisting of an initial flash, followed by an alternation of dark and light periods. All components but the initial flash have been called afterimages, or aftersensations. The initial flash has been called the sensation, or the primary sensation.

The logic of the experimenters, who named the sequence of events in this way, seems to have been that only the first event is the direct result of the stimulus. These experimenters possibly had initially expected the response to be as simple in its quantitative characteristics as the stimulus eliciting it. Therefore, all features of the response following the first component had to be accounted for by mechanisms other than those providing the initial reaction. This expectation with its implications represented only one of several alternative points of view, and there was nothing to confirm or deny the point of view in the visual phenomena themselves when they were discovered.

The nature of the experiential sequence following a photic pulse often provides a very complex set of events to be accounted for. The sequence poses the problem of what is to be assigned to photoreceptor processes and what is to be accounted for by the activity in the neuroretina, or other parts of the visual mechanism.

Recording of the neural impulses of the optic nerve by electrical instruments has furnished us with a means of determining a number of the characteristics of the processes that occur in the

eye. Problems such as just suggested regarding the determinants of brightness perception can often be answered by this procedure.

The Phenomenon. The feature of brightness perception selected for description in this chapter is one that was first isolated (Bartley) and later studied by electrical recording methods. The phenomenon is simply this: If a single photic pulse of proper

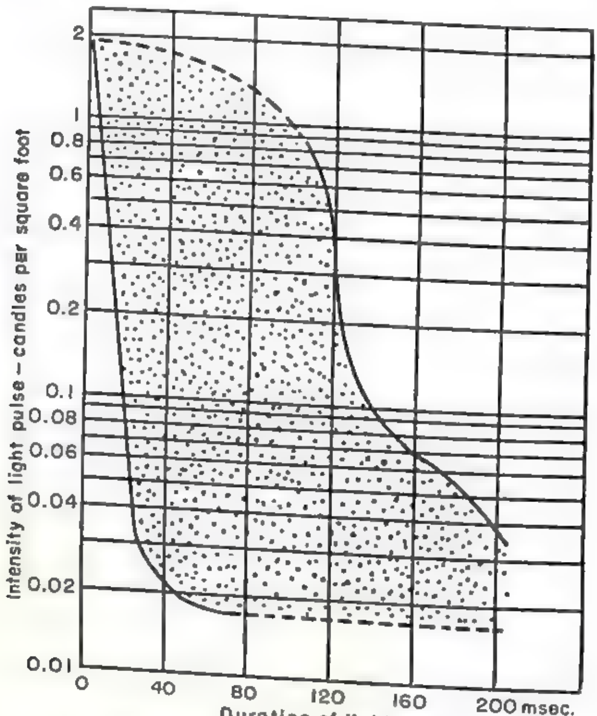


Fig. 75. Graph to show intensity-duration conditions under which a single light pulse gives rise to the two-flash experience.

intensity and duration is delivered to the eye, the sensory experience will be that of two nearly equal and separate flashes of light in quick succession. If the intensity of a pulse of light is quite feeble and its duration very short, the pulse will elicit but a single short flash. If the pulse is quite strong, or long, it will likewise elicit only a single flash. The two-flash experience occurs only when the light pulse is of an intermediate intensity and duration.

The Experimental Conditions (Perception Experiment). In the investigation of this phenomenon, a 5-in. square area of opal glass was viewed by the subject at a distance of 18 in. An Abney episcotister was used to vary the pulse duration, and a rheostat was used to vary luminous intensity. An Abney-type episcotister is one whose open sector, or sectors, can be varied in angular amounts during rotation. In the experiment, the episcotister was arranged so as to have only one open sector and to subtend only a few degrees at most. While the episcotister was kept in uniform motion, the subject could adjust the width of the opening from trial to trial (pulse to pulse). In this way the point at which the single pulse just produced the appearance of two flashes could be quite accurately obtained. Since it was not a matter of using a series of fixed flash durations, but rather a process of fine adjustments of duration from trial to trial, such measurements could be made quite readily. Figure 75 indicates the general range of the combination of duration and intensity throughout which the two-flash result was obtained. The ordinate of the graph is scaled in terms of logarithm of intensity in candles per square foot (c./ft.^{2}). The actual intensities used varied over a range of from 1.9 c./ft.^{2} down to about 0.02 c./ft.^{2} The duration varied from about less than 1.0 msec. to about 204 msec. In the columns in Table 16, for example, are indicated a few of the intensity-duration combinations that produce the two-flash experience.

TABLE 16

Intensity	Light pulses, msec.	
	Minimum duration to produce 2 flashes	Maximum duration to produce 2 flashes
1.90	Less than 1	
0.85	....	106
0.66	....	112
0.38	12.5	120
0.14	....	128
0.073	21	155
0.034	25	204
0.020	48	Off-response emerges

Figure 75 consists of two curves, the one to the left representing the boundary conditions at which intense short pulses become just intense and long enough to elicit the two-flash response. The curve to the right represents boundary conditions at which the two-flash response just disappears as intensity is further increased, and flash duration is further lengthened. All of the area on the graph bounded by the two curves represents the various combinations of intensity and duration that will elicit the

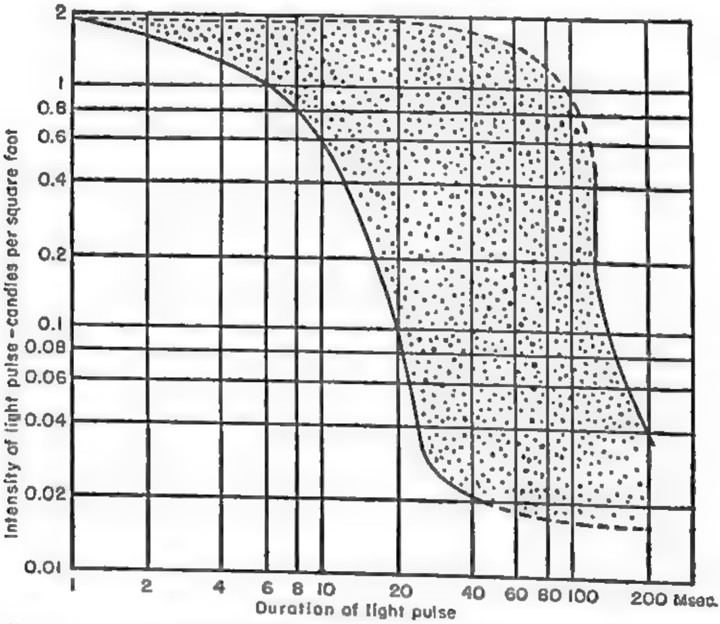


FIG. 76. The same as Fig. 75, except that both variables are plotted on a logarithmic scale.

two-flash effect. The graph indicates that intensity and duration are not related in simple reciprocity. Figure 76 was made from the same data as the preceding figure, except it is plotted on a log-log scale.

The second of the two flashes elicited by the single pulse of light is a quite unexpected feature, even in view of what is generally known about afterimages. Afterimages, or after-sensations, are generally thought of as being less impressive than the primary sensation itself. In this case, however, the second flash, at best, is about like the first and is uncomplicated by the

expected train of fellows (other phenomena which also could be called afterimages).[1]

In this phenomenon as just described, we have the problem of discovering whether or not there is anything in the record of the electrical activity of the optic nerve that would indicate whether the precursor of this second flash lies in the eye. Bartley and Bishop's earlier neurophysiological studies showed a very definite on-response in the optic-nerve record produced by a short pulse of light, and also an off-response in response to the termination of an extended exposure to light. Their mode of recording was used for the present analysis of the relation between light-pulse intensity and duration and the resulting optic-nerve discharge.

Apparatus and Procedure (Neurophysiological Experiment). The same apparatus for producing the light pulses was used in the present neurophysiological experiment on the optic-nerve discharge as for the perception experiment in observing the flashes.

The subjects in this experiment were lightly anesthetized rabbits. The anterior third of the animal's brain was removed so as to reach the optic nerves. Recording electrodes were placed on one of the nerves held from contact with other tissue. This suspension was necessary to avoid electrical shunting which prevented the obtaining of full-sized amplitudes in potential changes. The Abney episcotister with a single open sector was used for the light pulses. These were delivered 3 sec. apart. To begin with, the pulses in the series were very short and of low intensity. Pulse by pulse they were either lengthened or raised in intensity until a threshold-recorded response was obtained. From then on the variable of intensity or duration was further increased to obtain the pattern of optic-nerve discharge at each step.

The impulses from the nerve were made recordable by vacuum-tube amplification and made visible by a cathode-ray oscillograph. The face of the oscillograph tube was photographed at each trial by a camera specially built for the purpose. Sensitive

[1] Detailed studies of afterimages have generally disclosed that conditions producing them give rise not to a single pronounced afterimage but a series of them.

paper rather than negative film was used. This paper was 6 cm. wide and in long rolls that could be fed through the camera frame by frame or continuously at uniform speed. As you may know, the record on the cathode-ray oscillograph tube is a moving bright spot. When it moves fast, it is seen as a bright line. When it moves slowly, it is seen as a spot. The spot begins at the left-hand side of the tube-face (see Fig. 77) and moves horizontally across the tube-face. It retraces its path almost instantly and is ready for a new excursion every time a pulse of light is used as a stimulus for the experimental animal. The

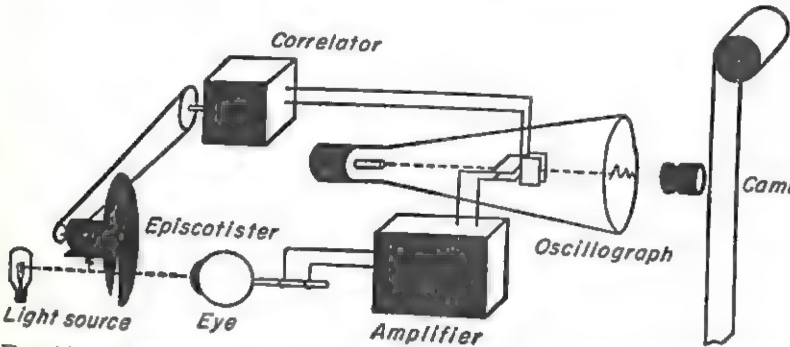


FIG. 77. A schematic diagram of the instruments used in stimulation and recording of the optic-nerve discharge.

electrical activity of the optic nerve causes an up-and-down motion of the spot. Consequently the record is a train of waves.

The camera was so designed that the wave activity could be photographed either as a series of transverse pictures across the strip of photographic paper or as a continuous line running the length of it. For present purposes, it was preferable to have a series of transverse records, each one being the optic-nerve response to a single light pulse.

Since the spot began in each excursion at the same place on the tube and traveled the same velocity across the tube, distances on the records could be translated into units of time. A series of transverse records, in effect, formed a column of records precisely aligned. The experimenter could, therefore, examine this column to determine the series of changes produced in the various parts

of the optic-nerve response, as intensity or duration of pulses were progressively increased or decreased.

Results. The threshold response consists in a small hump in the otherwise horizontal line in the record. This hump and the initial wave of the response to suprathreshold pulses is called the on-response. As intensity of pulse is increased above threshold, the on-response increases in amplitude and develops an irregular contour, generally presenting a somewhat bimodal appearance. Early in the step-by-step increase in intensity of pulse, the on-response is followed by an irregular train of wavelets. With further increase in stimulus intensity, the part of the discharge following the on-response quite soon changes. The wavelets are replaced by a single large wave. This grows somewhat in size as intensity of pulse is further increased. At maximum, this wave is almost as large as the on-response itself, although generally it falls somewhat short of equality with it. This wave is called the *secondary wave* and is labeled *S* in Fig. 78, which pictures the results just described. As light-pulse intensity is still further increased, the secondary wave quite quickly diminishes in size. In fact, it virtually disappears and is replaced by a train of wavelets. Were it not for the fact that the experimenter is able to scan a column of responses to pulses of increasingly great intensity, he could not be sure which of the irregular waves following the on-response is the remnant of the secondary wave. These have much the same appearance as those elicited by pulses too weak to give rise to the secondary wave.

From this description it would appear that the optic-nerve discharge contains the very component that we might expect as the basis for the second of the two flashes described in the early part of the chapter. The on-response is the nerve-impulse precursor for the first of the two flashes, and the secondary wave is the precursor for the second flash.

We must ask, however, whether the optic-nerve discharge exhibits a secondary wave throughout a middle range of *durations* as well as the middle range of pulse *intensities*. This we should expect it to do in order to compare with the sensory results obtained by varying pulse duration.

Manipulations of pulse duration very clearly demonstrate the emergence of a secondary wave as duration is increased and a disappearance of the wave as durations become great. As pulse durations are increased, the response to the light covers more time and manifests an off-response. Off-response begins by being very small, but as pulse duration is lengthened, it becomes quite prominent. Before the off-response becomes at all large,

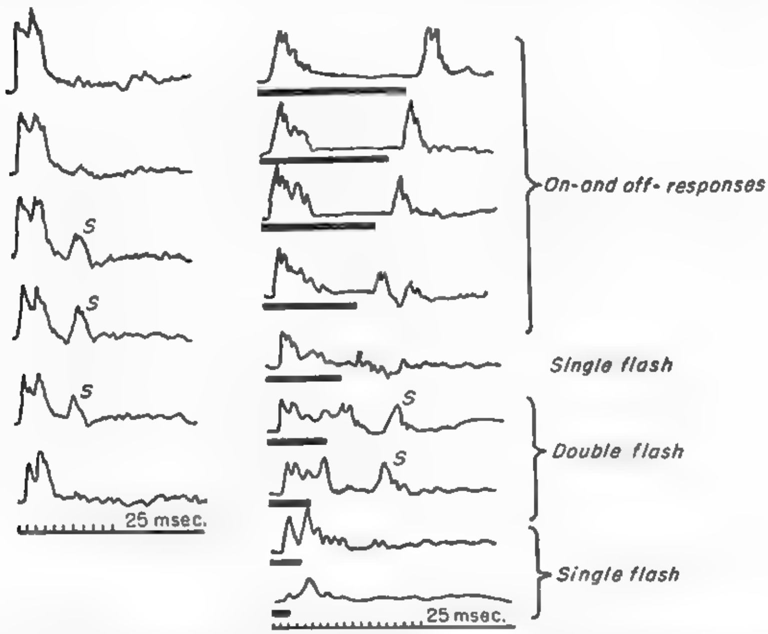


FIG. 78. The optic-nerve discharge in response to short pulses of light. The discharge includes a component S which is the precursor of the second flash with such stimuli. (*Bartley. By permission of J. exp. Psychol. and Amer. Psychol. Assn.*)

the secondary wave has become vestigial. But the off-response and the secondary wave can easily be distinguished from each other. It is clear from the records that the secondary wave moves to the right (becomes later) as pulse duration is extended. The secondary-wave remnant follows the off-response.

That it follows rather than precedes the off-response is very significant, for it makes sure for the first time that the secondary wave is actually an "aftereffect" and can be called an "after-

sensation," as the earlier workers wanted to call all responses except the initial sensation.

In this comparison of the optic-nerve discharge with sensation, we have an example of a very concrete relationship between a sensory event and a neural event that critically underlies it. That the comparison was not made on a single animal form (human) and perhaps in the same specimen does not obviate our confidence in the correlation. Neurologically the rabbit's eye is much the same as the human eye. The fact that the rabbit subjects were partially anesthetized is also no reason for expecting a marked difference in function, for anesthesia is a process affecting mainly the central nervous system rather than the peripheral. It is possible that were the animal able to forego an anesthetic, the secondary wave would have been somewhat larger than it was, hence more nearly like the on-response than it generally was found to be.

To say the least, the experiments just described form a very clear-cut example of success in correlating sensation with neurological conditions underlying it.

QUESTIONS

1. What are physiological mechanisms?
2. What might we call the brief light stimulus which gives rise to the experience of a flash?
3. What is the broad purpose of the present chapter?
4. What is the perceptual phenomenon that is under consideration?
5. What were the stimulus variables that were manipulated to provide for its appearance and disappearance?
6. What does the perceptual experiment have to do with after-images?
7. What is a cathode-ray oscillograph?
8. What is an optic-nerve discharge and how is it made observable?
9. Describe the optic-nerve discharge in terms of its components and their relation to the duration and intensity of the light pulse used.
10. How can comparisons between data obtained from human subjects and animal subjects be justified in this case?

CHAPTER 27

BRAIN WAVES IN CHILDREN AND ADULTS

Nerve impulses, as was indicated in the preceding chapter, manifest electrical effects. Accordingly, wherever masses of nervous tissue are found, it might be expected that some sort of electrical records could be obtained during tissue activity. Within the last two decades interest in recording effects from brain tissue has assumed wide proportions. Although many years ago it was discovered that electrical effects could be obtained from the brain, it was not until in the 1920's that interest in this sort of manifestation was renewed. This began after Berger had recorded electrical potentials from some patients who had cranial openings permitting placement of electrodes directly on or in the brain. Two sorts of interest developed in this country: (1) The laboratory study of brain activity in animals in which precisely controlled stimulation and recording were involved, and (2) studies upon humans in which recording from atop the skull was practiced. In the human studies less specificity and precision of stimulation and less detail in recording has been characteristic. The records from the human subjects soon came to be called electroencephalograms or simply EEG's. It was not long until the recording of EEG's became routine in certain hospitals. Although the greater number of studies of brain waves pertain to the human EEG, studies utilizing animals and cathode-ray oscillographs as recorders are better adapted to minutely telling what is happening in the brain.

It is our purpose to describe the kind of record obtained from the human. One of the very best investigations for this purpose is that of Lindsley, involving a comparative study of individuals of various ages. The younger subjects he used were 256 in number and ranged in age from 1 month to 16 years. One hundred

and forty were boys, and 116 were girls. His 75 adult subjects ranged from 17 to 64 years old, although most of them were 20 to 40 years old.

Apparatus and Procedure. The apparatus consisted of two independent amplifiers similar in over-all frequency-response characteristics, and supplied with filters so that high-frequency extraneous waves were eliminated. The recorder was a Westinghouse oscillograph, an instrument somewhat more sensitive than the usual electromagnetic inkwriter, owing to its avoidance of mechanical friction resulting when pens are used. It is an optical lever operated electromagnetically. The beam from a tiny mirror about 1 or 2 mm. long casts a point of light on moving photographic paper or film.

The electrodes used were small gold disks 8 mm. in diameter embedded in cupped bakelite blocks. The cups were filled with electrode jelly by which they made electrical contact with the scalp. These electrodes were held to the surface of the scalp by bandaging. The electrodes were placed over comparable regions of the occipital area in all subjects. The proper placement was ascertained by X-ray study and careful measurement. Two electrodes were used and placed 5 cm. apart in a line parallel to and 5 cm. lateral to the mid-line of the skull.

All subjects, except the infants, sat comfortably in a dark and sound-treated electrically shielded room. The infants were placed in a bed in the same room.

Results. Five types of waves were obtained. These, of course, include types that other investigators had already referred to by other designations.

The familiar alpha rhythm, originally described by Berger, and sometimes called the Berger rhythm, is *type I*. The waves of the alpha rhythm usually form one of the most prominent characteristics of the EEG. These waves range from about 8 to 13 per second. Except for children of three months or below, all subjects studied by Lindsley showed some evidence of an alpha rhythm. Whereas the frequency range just stated applied to the adults, the frequency in children was considerably lower. Figure 79 indicates the mean frequency of the alpha rhythm for each age group.

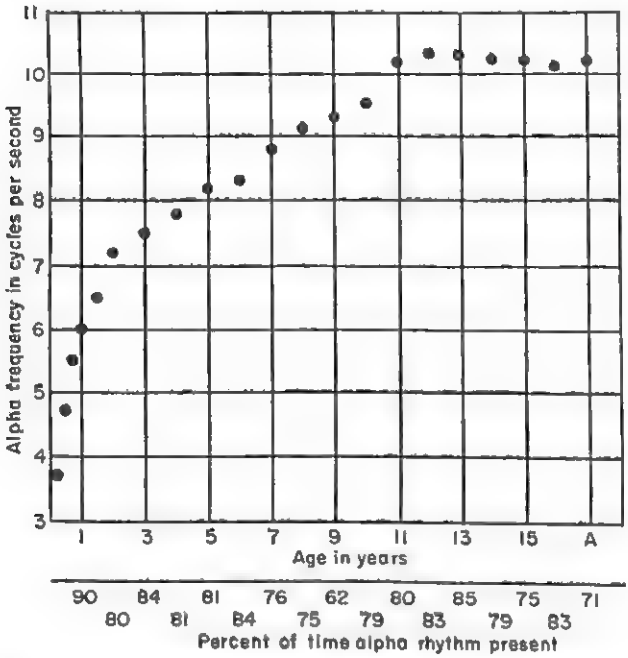


FIG. 79. Chart of the relation of the frequency of the alpha rhythm and human age. (*Data from Lindsley*, 1938. *J. gen. Psychol.*)

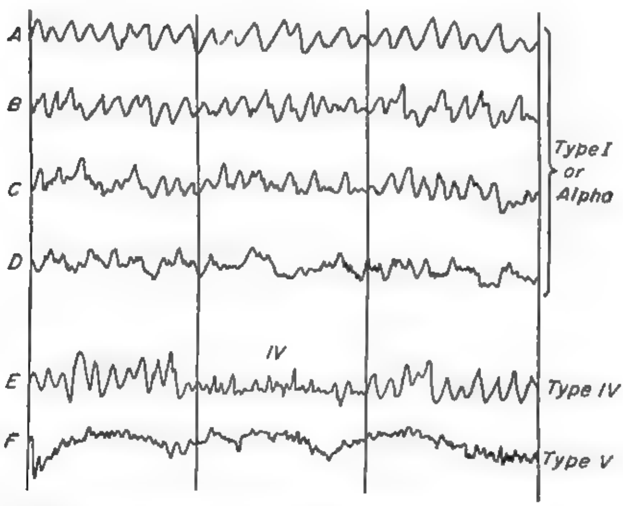


FIG. 80. Various types of brain waves found by Lindsley. (*Lindsley. J. gen. Psychol.*)

Figure 80 provides several representative records to show a range from EEG's with very simple and regular alpha waves to those that are much more complex and irregular in pattern.

Figure 81 shows a series of records obtained from a single child during its first three years of life. It will be noted that waves become more prominent as the very young infant increases in age. At four months very definite alpha waves appear. Typically the alpha potentials in younger individuals manifest greater amplitude than do the waves (potentials) in adult records. By

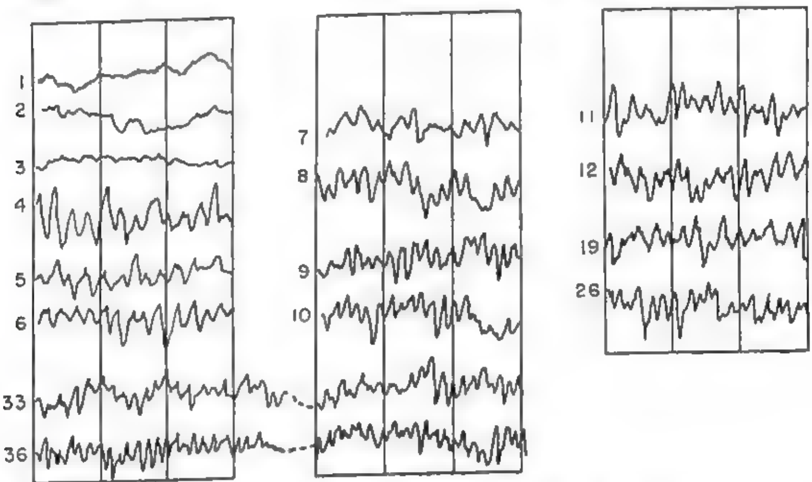


FIG. 81. Brain waves at different ages of a single child from 1 to 36 months. (*Lindsley. Ped. Sem. J. genet. Psychol.*)

the time adulthood is reached, the waves seldom exceed 40 microvolts. It was suggested that this difference might be due to the existence of thinner skulls in young people.

Of about the same height, namely, from 2 to 10 microvolts, are wave *types II and III*. Sensory stimulation and emotional state do not seem to alter them. Waves with frequencies between 20 and 50 cycles per sec. were identified by Berger and called beta waves. In adults, close inspection discloses that there are generally two distinct bands, (1) ranging from 20 to 30, and (2) from 40 to 60 waves per second. Of the two, the 20- to 30-waves-per-sec. band is the more prominent. Some authors have limited the term "beta waves" to this band. In the occipital regions the 40- to 60-per-sec. waves are more often seen than

those of the other band. Some authors believed that they had demonstrated the existence of a band of waves from 35 to 60 per second, and called them *gamma* waves. Lindsley applies the term *type II* to the lower band (20 to 30 per second) and *type III* to the upper band (40 to 60 per second). Both types II and III appear at lower frequencies in children than in adults. Furthermore, these two types have been observed in infants before the first appearance of alpha waves.

In many records, the alpha waves come and go instead of being continuous. When the alpha rhythm temporarily disappears, it is replaced by a train of waves of lower amplitude and higher frequency. These waves constitute what Lindsley called *type IV*. Waves of type IV are from one-quarter to one-half the amplitude of the preceding or subsequent alpha waves and are about double their frequency.

The appearance of type IV waves is not always simultaneous in both cerebral hemispheres. This lack of concurrence is evidence against the type IV waves representing merely a sort of interference between two alpha trains out of phase with each other. Type IV waves appear not only "spontaneously" but also as a result of whatever stimulation that succeeds in temporarily abolishing the alpha waves.

Waves of *type V* are found in the records of both children and adults. These waves are slow undulations with a period from about 0.6 to 1.5 sec. These waves, like the others, are longer in children than in adults. The amplitude of type V waves is generally from one-half to three-quarters that of the alpha waves. Type V waves seem to be unaffected by stimulation.

Other Waves. Since Lindsley, in this study, dealt only with waking subjects and those not suspected of brain lesions such as tumors or those known to be epileptic, he did not obtain all the kinds of records described in the EEG literature. During sleep the alpha undulations in the record are replaced by slower waves that have been called *delta* waves. Delta waves vary in frequency from 0.5 to 5 per second. They are sometimes spoken of as being 2 to 3 per second.

Lindsley did find, however, in about 15 per cent of his subjects, occasional evidence of a wave frequency of 2 to 4 per second in

children (1 to 5 years old), and 4 to 7 per second in older children and adults. In two subjects (children), a slow frequency predominated in their EEG's. Since these slow waves were found in waking individuals, Lindsley refrained from identifying them as delta waves, even though they fell within the delta frequency range.

Lindsley plotted the increase in frequency of wave types I, IV, and V with age and found the curves to resemble the curve of increase in brain weight as a function of age. Correspondence was absent when increases in wave frequency with age were compared with rate of growth of gross body structures.

Since various claims had been made by certain authors regarding correlations between the wave patterns found in EEG's and personality traits, etc., Lindsley made certain tests bearing on this matter. He found no significant relationships between EEG characteristics and the personality aspects that he tested. For example, he used the aspect of intelligence as determined by the Stanford-Binet and Otis tests, dominance or assertiveness as indicated by the Allport Ascendance-Submission Test and emotional stability as measured by the Woodworth-Mathews Personal Data Sheet. Information on these three matters was available for 88 of his subjects ranging from 10 to 14 years of age. The information was compared with the percentage of the time the alpha rhythm was present in the EEG for each of these subjects. The correlation coefficients were almost zero for each of the three personality characteristics. The same failure resulted in the attempt to use the factors of regularity and uniformity of EEG in the correlation procedure.

Lindsley employed some of the usual procedures of stimulus variation to test their effects upon the brain potentials. Tactile, visual, and auditory stimulations were used. Visual stimulation in the form of isolated brief pulses of light was found to be about the most effective for the study of time required to produce a change in the alpha rhythm and the time required for this change to disappear. Sensory stimuli momentarily abolish, or greatly diminish, the amplitude of alpha waves. This has conventionally been called *blocking*, although the implications involved in the term are not necessarily appropriate to the situation. The

time elapsing until the reappearance of the alpha waves was called *recovery time.*

To study blocking and recovery times, pulses of light requiring about 1/8 sec. to reach full intensity and about 1/5 to 2/5 sec. long were used. A definite tendency for blocking time to decrease with age was found. It was suggested that this might be related to the decrease in the duration of the alpha waves as age increases. Recovery time manifested no consistent relation with increase in age.

Conclusions. This study of brain potentials, like many others, was not conducted to test a formally expressed hypothesis. A number of studies described in the various chapters up to now have had this characteristic in common with the present study. In their outward aspects, therefore, they have not seemed to exemplify the rules laid down in the early chapters on the nature of experimentation.

If we look into the present study, we can find the essentials of a good experiment as described in the early chapters. We can find certain presuppositions and implicates upon which the exploration was founded and which it did much to test. Your recognition of these assumptions will do much to transform what seems to be sheer exploration into obvious experimentation conforming to the pattern that we earlier described.

In Lindsley's study of human brain potentials, you have been given the findings of a basic and dependable investigation in this area. It must be pointed out, on account of the prevalence of present-day interest in brain potentials and enthusiasm in interpreting them, that a very special background is required for doing dependable research in this field. It must not be thought that the mere ability to run off yards or even miles of EEG records and make certain correlations between some perceived characteristic in them and some alleged personality trait is, in general, dependable science, or, in particular, the furtherance of neurophysiology or psychology.

No one but an individual fairly well versed in neurophysiology is in a position to assign meaning to brain potentials. The naïve tend to entertain many kinds of unwarranted expectations in regard to what brain potentials indicate. The untrained almost invariably expect to find relations between specific aspects

of the EEG and all kinds of personality traits. In fact the literature contains assertions that some such correlations have already been demonstrated.

We need to know a great deal more about the functional relations between the pathways in the brain and which group of elements are producing the components of the EEG before we can go very far in understanding what they mean. For example, we need to know enough about the physiology of various cortical pathways and centers to be able to say whether shifts in EEG pattern represent shifts in activity from one cell group to another or merely changes in function in the cell groups initially recorded. As it is now, it has been taken for granted by many people that epileptoid charges in the EEG represent modifications of the alpha rhythm. This is true despite the evidence from neurophysiology to indicate that in such conditions the alpha rhythm disappears, and the spikelike activity replacing it is the manifestation of the activity of certain other groups of neurons.

For purposes of providing insight into the way that the brain potentials are studied in the neurophysiology laboratory, and to describe certain basic findings, the following chapter continues the topic of brain potentials.

QUESTIONS

1. Tell what is meant by "brain waves," in a way to be understood by one whom you suppose never heard of them.
2. Contrast the two sorts of interests that have developed in this country with reference to the recordable electrical activity of the brain.
3. What kind of apparatus is needed for recording brain waves?
4. How may brain waves be classified according to Lindsley?
5. How do the EEG's of young children and adults differ?
6. What is meant by "blocking"?
7. What cautions or limitations ought to be recognized by psychologists in their interests in brain waves?
8. Was Lindsley attempting to test some hypothesis in his study of brain potentials? If not, explain what he was doing.
9. Could a correlation between the finer details of the EEG and traits of personality, or personality types, be expected? Justify your answer.
10. What feature of the EEG has received most attention? Why do you suppose this is so?

CHAPTER 28

BRAIN WAVES AND CORTICAL RESPONSE TO SENSORY STIMULATION

In the preceding chapter the recordable electrical potentials of the human brain (brain waves) were described particularly on account of the widespread interest in such phenomena in psychological circles. In the chapter's conclusion, the suggestion was made that you ask yourself what sorts of assumptions are involved in the study of brain potentials. What assumptions do you have when you concern yourself with brain activity? Could not certain viewpoints be entertained in psychology that would preclude attention to this kind of subject matter?

It is fitting, in light of present-day interest in brain potentials, that experimental psychology students obtain a view of brain potentials from the neurophysiologist's standpoint. In neurophysiology it is customary to use animals for extensive analytical work and to provide precise forms of stimulation and to use the most sensitive and faithful forms of amplification and recording available.

In such studies the attempt is made to elicit specific effects by use of specific forms of stimulation. For example, when using the visual pathway as a sensory avenue for stimulation, it is customary to use short pulses of light that are controlled in intensity and duration. The behavior of specific groups of neural units, or neural pathways, can be disclosed, based upon (1) precise manipulations of stimulus variables, (2) known recording sites, (3) character of specific responses to such stimulation, and (4) animal states.

To provide the most faithful recording of detail, the cathode-ray oscillograph must be employed as the recording device.

For our purposes, the kind of investigation made by Bartley,

O'Leary, and Bishop will exemplify the technique used in the analysis of the components of brain-wave records. These records will show the relation of some of these components to the appearance and disappearance of alpha waves.

Procedure. The above-mentioned authors used a kind of strychnine technique earlier developed by Dusser de Barenne. The application of a very small amount of dilute solution of strychnine to a 1- or 2-mm. spot on the cortex induces changes in the size and other characteristics of the spontaneous cortical activity. In the present study an analysis of the cortical response was accomplished by means of the strychnine-induced changes that it manifested. Bishop and Bartley had earlier studied the specific response of the unstrychninized optic cortex to brief electrical stimuli applied to the stub of the optic nerve in the rabbit. The animals were prepared in much the same way as was described earlier for recording the activity of the optic nerve, except that the cortex was left intact. The use of ether anesthesia was found preferable to a number of others for the purpose of this investigation. The surface of the rabbit brain is smooth (not convoluted), and for this reason the rabbit brain is convenient to use. It is also convenient because neurologically each eye is almost completely represented in a single hemisphere.

The authors in the present experiment used needle electrodes that could be made to penetrate the cortex to considerable depths or merely to lie under the surface. For purposes of cortical exploration and recording, the posterior one-half to three-quarters of the skull over one hemisphere was removed. The dura mater[1] on that side was also removed.

The first step in experimentation was to determine points on the occipital cortex at which the most satisfactory cortical responses could be obtained. When these were once found, records were obtained to represent the unstrychninized condition. After this, strychnine was applied to an area a few centimeters from the recording site. This was done by means of a tiny wisp of cotton twisted to a point and dipped into a 1 per cent solution of strych-

[1] The dura mater is the tissue sac which encloses the brain. It contains cerebrospinal fluid that fills the space between it and the brain.

nine sulphate. The cortical surface was freed from excess fluid, and the damp wisp was touched to it. In this way 1 or 2 sq. mm. could be treated with strychnine. Further recording was reinstituted to note progressive effects. Depending upon the distance from the recording site to the strychninized area, the effects showed up immediately or after a period of seconds, or even minutes. In some cases, the recording area itself was directly strychninized. The effects in the record, once they became apparent, manifested greater and greater involvement. The strychnine tended to spread to some degree along the moist surface of the brain. In this way, as well as by penetration of the tissues themselves, the strychnine was carried eventually to surrounding regions. The effects manifested at the recording sites were mild enough at first to permit observation of slight involvement. The involvement progressed slowly enough to permit recording a number of stages. Some of the effects were at times reversible by washing with Ringer's solution (a physiological salt solution).

The Activities of the Normal and Strychninized Cortices. In an anesthetized animal a specific response in the occipital (visual) cortex of the brain can be elicited either by exposure of the eye to a pulse of light or by a short electric shock to the stump of the optic nerve—the eye having been removed. A description of the response in the rabbit or in the cat follows.

The response is a complex wave pattern, as pictured in Fig. 82. It consists of three components, one of which is at least partially made up of impulses *arriving* at the cortex. The other two components represent activity occurring *in* the cortex. These latter two components are labeled *B* and *C* in the uppermost diagram in the figure. *B* is a diphasic wave that with the usual film or paper speeds used in recording is rather abrupt and sharp. The other component *C* is a triphasic wave with a period about ten times as long as that of *B*. *B* and *C* are partially concurrent, as is suggested in the diagram. Therefore certain features of the wave picture represent their algebraic summation.

A single pulse of light, for example, sets off this complex of events called the cortical response. When the cortex of the animal is "spontaneously" active in such a way that an irregular pattern of waves and wavelets shows up in the record, the cortical

response may be partially obscured. It can, however, generally be identified. Occasionally the spontaneous record is of low amplitude, thus enabling clear observations of what the cortical response looks like.

Often the specific cortical response does not end with the undulations pictured in the upper diagram, but instead seems to constitute, in addition, a train of several waves of the same frequency as component *C*. In the rabbit, the waves in this train commonly

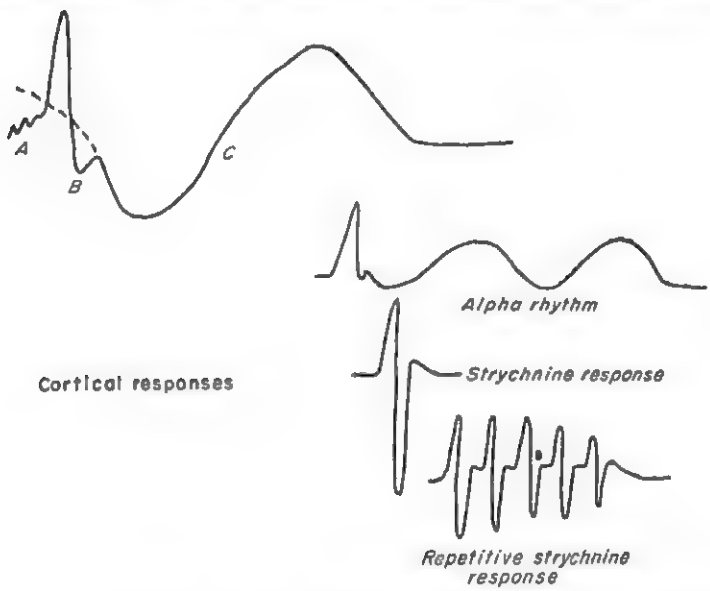


FIG. 82. The normal cortical response to a pulse of light or an experimental electrical stimulus, and also the strychnine response.

have a frequency of about 5 per second. This is the rate of the dominant waves in the spontaneous record, which we are at liberty to call the animal's alpha rhythm. This elicited train of waves can best be seen when the record is of very low amplitude or when the spontaneous alpha waves are absent. Thus, single pulses of light elicit a cortical response that may either be brief or may include a train of alpha waves.

It was also found that the amplitude of the *B* component of the cortical response varied in height depending upon its temporal relation to the spontaneous alpha waves. In some phases of the alpha cycle, the response was of low amplitude; in others, high;

and in still others, of medium height. It was therefore possible to "tune" repetitive stimulus pulses to the intrinsic rate of the alpha rhythm. The tuning was made, at times, to relate to the phase of the cycle in which low-amplitude responses were produced; and at other times, the tuning was made to relate to other phases in which medium or high amplitudes were produced.

When the light pulses were intense, they often produced a result different from that represented in tuning. The process involved has been called "driving." Driving is the alteration of the intrinsic cortical (alpha) rhythm so as to coincide with the repetitive rate of the stimulus series. Driving occurs best when the stimulus rate and the intrinsic cortical rate are the same. Driving then consists simply in shifting the waves in the cortical train to coincide with the stimulus series.

When the repetitive rate of the pulses in the stimulus train is increased, driving soon may be obliterated. But if the stimulus rate is some small multiple (twice as great, for example) of the rate of the alpha rhythm, the rhythm may again be driven. Apparently the driving process may, in some animals, proceed quite successfully up the scale from the point at which the stimulus and the alpha rates are equal to the point at which the stimulus is several times the intrinsic alpha rate.

The phenomena of tuning and driving indicate the existence of a close connection between the immediate activity of incoming peripheral impulses as they reach the cortex and the intrinsic ongoing spontaneous activity as expressed in the alpha rhythm.

Another way of analyzing cortical behavior is by using paired stimuli (two pulses of light or two electric shocks). By using different time intervals between stimuli, the response to the second stimulus can be studied. Bartley found that if the second of two stimuli was applied soon after the first, no demonstrable cortical response was discernible in the record. As separations were made greater, a small response first appeared and then began to increase in size until a separation interval was reached at which the second response was identical to the first. The interval just permitting this was of the same order as the period of the alpha rhythm. Figure 83 indicates some typical results from using the paired stimulus method. It will be noted from

the records that two brief pulses of light 70 msec. apart do not provide for two responses. The second of two light pulses 90 msec. apart possibly elicits a just suprathreshold response. The second of two pulses of light 137 msec. apart elicits a response. By the time that the separation is 196 msec. apart, a nearly if not full-sized response is elicited by the second pulse of light.

In connection with the separation of the two stimuli is the matter of relation to the natural alpha frequency. The separations of 90 msec. and 196 msec. are approximately one-half and

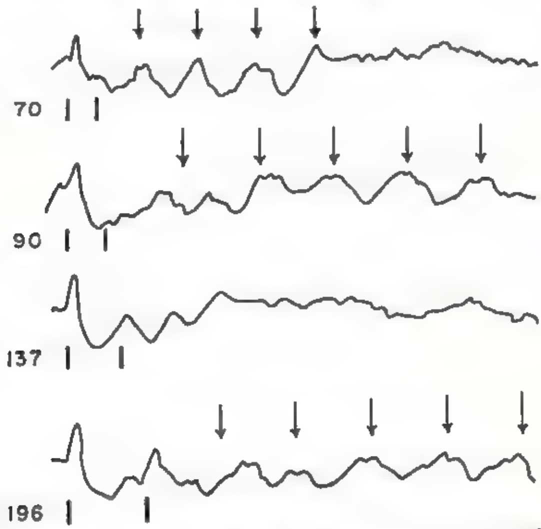


FIG. 83. The cortical response (rabbit) to paired stimuli (light pulses) variously spaced. (*Bartley. Vision. New York: Van Nostrand.*)

full alpha periods, respectively. It will be seen in the records that paired stimuli in those cases result in not only the initial spikelike responses but also in a train of alpha waves. Naturally in both cases the stimuli were administered at a random instant and thus without regard to timing of the natural alpha activity spontaneously in existence. On this account, one of two results would be expected to occur. Either the stimuli would not be able to set up their own alpha train, owing to the asynchrony of the stimuli with the spontaneous alpha waves, or else they would succeed in driving the cortex. If the latter happens, it may take a cycle or two to get the train fully shifted from the spontaneous to the new timing of the alpha waves. This shifting

is evident in the records for the 90- and 196-msec. pairs of light pulses. The downward-pointing arrows in Fig. 83 indicate the timing of the alpha waves thus set up.

The stimuli separated by 137 msec. are fully out of timing with the spontaneous alpha activity and the intrinsic periodicity of the cortex, hence all that can happen is a response to the first stimulus and a fractional response to the second stimulus. The response to the second stimulus is supposedly made possible by recovery of part of the elements (not single neurons) by the time the second stimulus was administered.

An analysis of the relation between cortical response and alpha activity by the strychnine technique yielded strong evidence that two distinct sets of neural elements are responsible for the *B* and *C* components of the cortical response. It also indicated that the elements responsible for the *C* component and those that produce the alpha rhythm in the spontaneous record are identical.

The application of strychnine tended to alter the cortical activity. One of the first effects was to diminish the amplitude of the alpha waves. Ultimately it obliterated them entirely. Paralleling the diminution and obliteration of the alpha waves was the diminution and complete disappearance of the *C* component. At the same time the *B* component of the cortical response was greatly augmented. The response at this time became all-or-none. Sometimes at this stage several stimuli had to be given before any response would occur. Furthermore, spontaneous waves having a form identical with the response waves would occur at random intervals. At one stage in strychninization a single stimulus would set off trains of response waves, each wave apparently involving the same neural elements as its predecessor. A typical train is shown in Fig. 82. Finally, when strychninization became extreme, all wave activity vanished, and no response was elicitable.

The spikes constituting the cortical responses in the strychninized cortex are augmented *B* components. Thus convulsive activity in the cerebral cortex represents the activity of those elements responsible for the *B* component.

Cortical convulsions produced experimentally, as by strych-

nine, and those in cases of epilepsy probably occupy the same neural elements. If this is the case, we may suppose that peculiar spikelike EEG of the epileptic seizure represents activity of cells not predominantly responsible for the alpha rhythm. This is in contradiction to the conventional assumption that these spikes are a distorted kind of alpha wave. In the epileptic record one cannot see concurrently an alpha rhythm and the repetitive spikes. So it has been deduced that the usual alpha waves have simply shifted into the spikelike waves.

Conclusions. The aim of this chapter has been to indicate how, with precise experimentation involving controlled peripheral inputs, the pattern of cortical activity can be analyzed. In addition, the aim has been to suggest that much more of this kind of analysis is necessary before we can understand what is happening in the cortex.

QUESTIONS

1. What is the objective of the present chapter? How does it differ from that of the preceding chapter?
2. What is the strychnine technique, and how was it used as a tool in the investigation described in the present chapter?
3. If EEG's can be recorded from the human scalp, why are animal studies such as that of Bartley and colleagues necessary?
4. Why was the rabbit the animal form used?
5. What are the components of a cortical response to brief optic nerve or retinal stimulation?
6. What are strychnine spikes?
7. What are some of the facts relating one component of the cortical response to alpha activity?
8. What is meant by "driving" the cortex?
9. What is meant by "tuning" in studying the behavior of the optic cortex?
10. What is illustrated in Fig. 83?

CHAPTER 29

LEARNING RATE IN A SINGLE SYNAPSE

Learning curves are generally descriptive of changes in the overt performance of human or animal subjects. Learning has generally been thought of as change in performance mediated by the central nervous system. The learning function has varied in keeping with the type of performance measured and has ranged from a negatively accelerated to a positively accelerated curve. In a few instances, straight-line changes have occurred. Shurrager and Shurrager assert that these curves should be spoken of as learned-response curves rather than learning curves, since they are representative only of a by-product of internal or central learning. In keeping with this idea, the authors assert that a *true learning curve* should be an expression of the rate at which some change progresses within the central nervous system. Be this as it may, the idea of measuring, as directly as possible, the changes that occur at given loci in the central nervous system is a very enticing expectation. The attempt to make such measurements would be a very important undertaking. Shurrager and Shurrager recently reported this type of study. They described what they believed to be learning in a single synapse.

It is our intention to describe their investigation. P. S. Shurrager, some years ago, produced conditioned responses (CR) in spinal preparations. These were called spinal CR. The original curve of spinal conditioning was negatively accelerated, when plotted to show the relation between percentage of trials and percentage of learned responses. The first 40 per cent of the trials exhibited the first 50 per cent of the conditioning.

What Shurrager and Shurrager actually did in the present investigation was to measure the rate of modification of a single nerve cell's ability to respond. This was taken to be its learning.

Materials and Technique. The spinal cord of the experimental animal (dog) was severed at the third lumbar spinal root, leaving the essential blood supply intact. Both semitendinosus[1] muscles were exposed and excised. One electrode was attached to the tip of the tail and another to one hind foot. Two quite weak stimuli (electric) were applied to the tip of the tail. The second of these was accompanied by a shock to the paw sufficient to produce a full contraction of the entire semitendinosus muscle. The stimulus to the tip of the tail was the conditioned stimulus (CS), and the stimulus to the paw was the unconditioned stimulus (UCS). After the learned response of the initially responding motor unit[2] reached 100 per cent, *i.e.*, when the conditioned stimulus (CS) elicited a response ten times in a series of 10 trials, each succeeding series of 10 trials was taken as a unit for comparison. Extinction was begun by using shocks to the tail (CS) without the accompanying reinforcing shocks (UCS) to the paw (see Fig. 84). This extinction process was kept up until the muscle discontinued to respond to any stimulus in a unit series of 10 trials. The full series of events in the experiment represented the following pattern: conditioning, extinction, reconditioning, reextinction, through the fourth conditioning and the fourth extinction. Vincent curves[3] were plotted from the results. This was done in order to determine whether conditioning and extinction changed rate as evidenced by the number of responses elicited in each unit of conditioning (10 trials) and each unit of extinction (10 trials).

Results. The curves in Fig. 85 represent the velocity of the successive conditioning series. For the next procedure in treating the data, the initial portions of the negatively accelerated curves used were selected by observation, the remaining portions beyond which the curve altered sharply in slope were discarded.

[1] This is a spindle-shaped muscle of the posterior and inner part of the thigh.

[2] A motor unit is an efferent nerve fiber and all the muscle fibers it activates through connection between them and its fibrils.

[3] Vincent curves are learning curves in which the data in one learning series are plotted with others by a method in which all series are reduced to the same number of steps.

The straight lines fitting the data were derived by applying the method of least squares.[1] These lines possess different slopes, the steeper the slope, the higher the rate represented. Such slopes are often spoken of as representing a function plotted

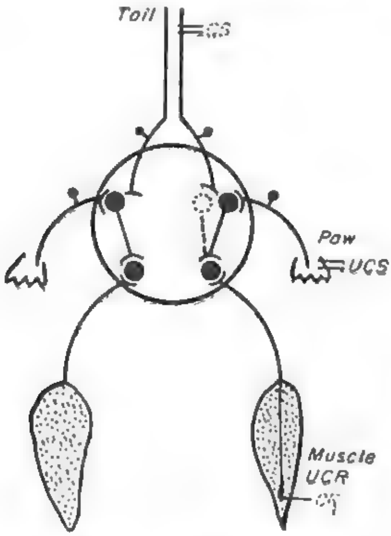


FIG. 84. A schematic diagram of a cross section of the isolated spinal cord, indicating the essential neural aspects of localization of the learning function. CS is a weak stimulus to the tip of the tail. UCS is stimulation of the paw producing UCR, a contraction of the entire muscle. CR is a response of an all-or-none (none-or-all—Shurrager) unit group of muscle fibers innervated by a single motoneuron (motor unit). The dotted neuron represents the processes which result in the learned response of the individual motoneuron as it becomes able (learns) to initiate its nerve impulse in response to a specific sensory conditioned stimulus. (*Shurrager and Shurrager. By permission of J. exp. Psychol. and Amer. Psychol. Assn.*)

against time or against numbers of repetitions. Thus the slopes can be stated as a ratio or fraction, and a number of them can be compared with each other. In this case they were spoken of as dc/dt slopes. The dc/dt slopes were plotted against the four

[1] An accepted method in mathematics (statistics) of deriving from a group of carefully made yet slightly irregular observations of a phenomenon, the most probable value of its unknown quantities, as for example, the position of a straight-line curve that would best represent a string of points not quite lying in a straight line. It is a kind of complex "averaging" process. For further understanding consult a statistics book.

conditioning series, as indicated in Fig. 86. The curves formed by these four points represent the velocity of the velocity or the internal acceleration curve of cumulating ability to respond. This the authors take to represent the rate of learning at the single synapse.

The following description of the situation and the deductions made from it are intended to clarify why the curve in Fig. 86 is to be taken to represent what Shurrager and Shurrager say it does.

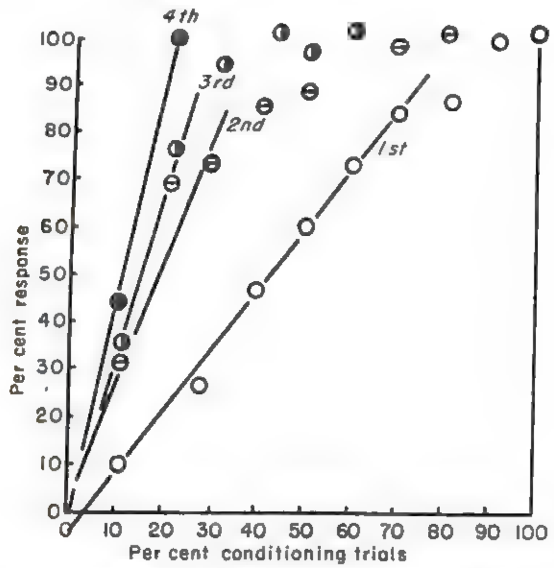


FIG. 85. Curves to show the progressively increasing velocities of the successive conditioning series of the response of the motor unit in the isolated spinal cord—muscle preparation. (*Shurrager and Shurrager. By permission of J. exp. Psychol. and Amer. Psychol. Assn.*)

The conditioned response (CR) is first manifested as the contraction of a small group of fibers in one end of the semitendinosus muscle. This group of fibers (1) acts in an all-or-nothing manner (as a unit). Sometimes this all-or-nothing unit can be conditioned and extinguished before a second unit starts to respond.

With the continuation of conditioning trials additional all-or-nothing muscle units (motor units) may become conditioned as is indicated by their participation in the contraction. These additional units (2, 3, 4) may be those toward the other end of

the muscle. During the over-all process of conditioning the initial unit (once fully conditioned) always continued to respond. This "superior" motor unit (since it was first) was found to learn (become conditioned) more rapidly in each successive attempt at reconditioning. Likewise extinction occurred more quickly from period to period. Some preparations (tissues, etc., made ready for experimenting upon) were run through as many as 19 periods of conditioning and extinction.

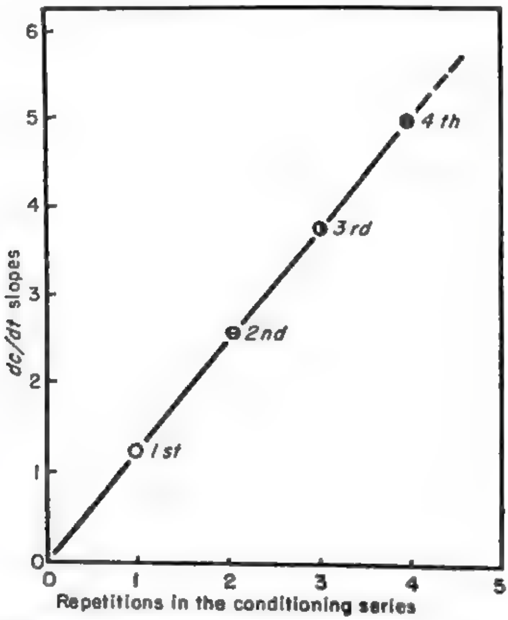


FIG. 86. The relation between the *dc/dt* slopes and the repetitions of conditioning of a single motor unit (a single synapse involved). The term *dc/dt* represents the ratio of increments of conditioning to increments of trials. (*Shurrager and Shurrager. By permission of J. exp. Psychol. and Amer. Psychol. Assn.*)

Even when extinction of the conditioned response was complete and no visible response to the condition stimulus could be obtained, some sort of a residue remained. This residue, or memory, was demonstrated by the increasing rate at which the conditioned response could be reestablished. This effect was cumulative, as was shown in the increasingly fewer trials needed to recondition a motor unit, and was susceptible to measurement. Measurement consisted in (1) ascertaining the *dc/dt* slopes (ve-

locities) of the response curves for the successive periods of the development of response (learning), and (2) plotting each of the slopes of the curves against the serial number in the reconditioning procedure that it represented, as in Fig. 85. The curve formed by these points was called the true-learning curve of a single motor synapse of the all-or-none (called none-or-all by Shurrager and Shurrager) muscle unit. Figure 85 was the outcome from five animal preparations.

Although the phenomena measured turned out to exemplify the straight-line function just described, the question of what was being measured still remained to be answered. The following discussion is intended to do this.

Further Discussion. The very first observable response that was set up in the semitendinosus muscle was a contraction of a small cluster of muscle fibers. This cluster was such as to be an example of muscle fibers included in what is known as a motor unit. It therefore must have been activated by a single motor neuron. On this account, the Shurragers concluded that the localized site of learning was the motor-neuron synapse. Had the learning taken place in a wider field, the initial manifestation would have been a contraction including much more of the muscle.

For example, the sensory impulses arriving at the internuncial neurons impinge on the dendrites and cell bodies of many neurons in the "pool." Were the learning to have occurred either at the sensory or internuncial synapses, a number of motor neurons would have been involved and the initial contraction in the conditioning experiment would have involved many more fibers than were actually involved.

Concluding Remarks. This study has demonstrated something about the fundamental changes that take place in the individual synapses or neurons. By virtue of the fact that not all neurons learn equally easily, a single superior neuron learning first made it possible for the Shurragers to single out its behavior and show that its learning curve was a straight line. While the study of the behavior of single neurons and their connections is not strictly psychology, it is good physiology and represents the kind of information of which we ought to have much more. It

is the kind of information upon which a much better understanding of what psychologists call learning could be built.

QUESTIONS

1. In what very essential respect did Shurrager's and Shurrager's study of learning differ from the customary studies in learning?
2. What do these investigators say the true learning curve should represent?
3. What is negatively accelerated learning?
4. Give the essential details describing the tissue preparation used in the Shurrager and Shurrager learning experiment.
5. Describe the experimental procedures used to test whether learning could occur.
6. How does the "motor unit" enter into the present investigation?
7. Why was the response repeatedly conditioned and extinguished?
8. What fact was taken as evidence of learning?
9. What was the shape of the learning curve in the Shurrager and Shurrager investigation?
10. Why was this learning considered to be that of a single synapse?

PART VI

COMPARATIVE PSYCHOLOGY

CHAPTER 30

DIRECTIONAL ORIENTATION IN MAZES

One of the most frequent observations that has been made in connection with the use of animals in mazes in learning experiments is that they seem to develop a general spatial orientation with reference to the maze or problem box used. Dashiell became interested in this problem and set out to see what could be done to solve it. Since the problem is a basic one in the understanding of animal behavior, the present chapter is to be devoted to it. For this purpose we shall review the experiments that Dashiell conducted. He performed a number of experiments involving a variety of mazes. The first three experiments were preliminary, each employing an essentially different type of maze. Each experiment was in a particular way capable of demonstrating the animal's ability to orientate itself.

The essential problem, as Dashiell saw it, was to ascertain what orientational characteristics of behavior the animals actually show in a variety of situations, what these situations have in common, what the actual variables are, and what the constants are. This involved primarily an attempt to see what the animals *could* utilize for purposes of orientation.

In the *first experiment*, a maze was used whose partitions could be changed from trial to trial, but whose direction from starting point to food box was essentially constant. The position and orientation of the maze in the experimental room was kept constant, the starting point was always on the same side, and the food box was always on the side opposite to the starting point. Figure 87 is a diagram showing 25 different variations of the partitions in the maze. Despite the constant changes in partitions, the animals made progressively better runs. That is, they reduced their errors from day to day.

There were two kinds of errors possible: (1) entrance into blind alleys or compartments facing toward the food box, and (2) entrance into those facing in the opposite direction. The animals did not make an equal number of the two kinds of errors. Throughout this chapter, we shall call the first kind of errors forward or *F* errors, and the second kind backward or *B* errors.

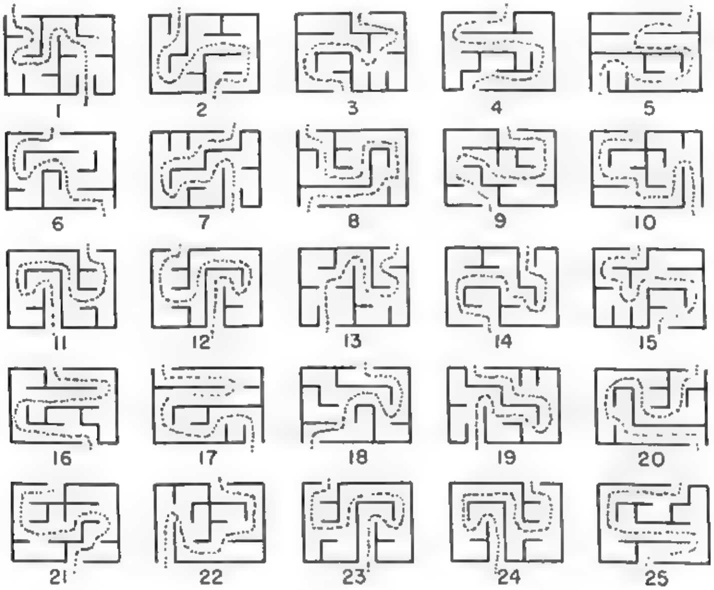


FIG. 87. One of the mazes used by Dashiell, in which the partitions could be rearranged in 25 different ways, making the true path different in each case. (*Dashiell. Comp. Psychol. Monogr., Johns Hopkins Press.*)

Figure 88 indicates the number of errors of each kind made. The upper half of the graph indicates errors toward the food box; the lower half, errors away from the food box. It will be noted that of the errors made those toward the food box (*F* errors) remained more or less the same as trials went on, but those away from the food box (*B* errors) were gradually and materially reduced. Furthermore, there were, in general, more errors made in the direction toward the food box.

The *second experiment* involved the use of what Dashiell calls the *F–B* maze. It is pictured in Fig. 89 (left). It, too, contained forward and backward blind alleys. At the end of the maze proper, the correct pathway was divided. By use of doors

at x and y, the pathway could be made to lead to a food box located either at f or b, as the experimenter chose. It is obvious that by use of this type of maze the location of the food box (goal) could be radically changed in position. Part of the time,

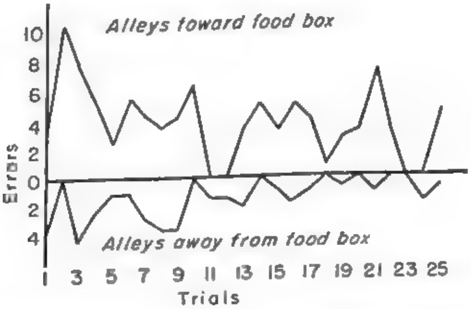


FIG. 88. Graph showing the number of errors (entering alleys) toward the food box and away from it as trials progressed. (*Dashiell. Comp. Psychol. Monogr., Johns Hopkins Press.*)

or for one set of rats, it could be placed at b, close to the entrance or starting point. At another time, for the same rats, or for a second set of rats, it could be in the opposite direction, at the end of the maze opposite the starting point. The forward alleys are designated F, and the backward alleys, B. It will be further

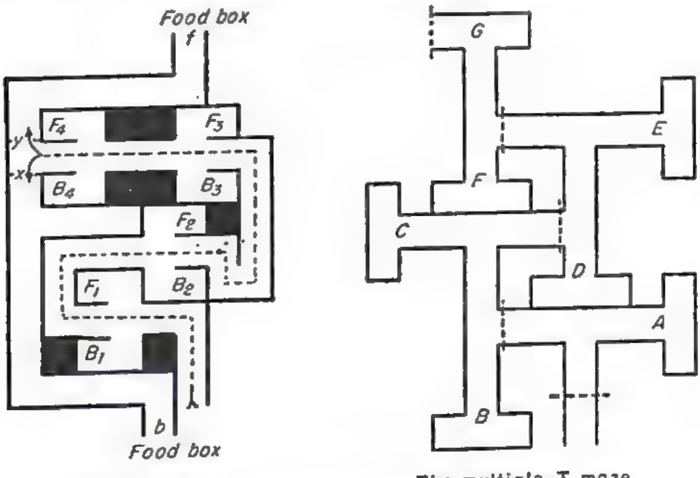


FIG. 89. Two kinds of mazes (left and right) used in studying spatial orientation. (*Dashiell. Comp. Psychol. Monogr., Johns Hopkins Press.*)

noted that the *F* and *B* alleys lie opposite each other along the true pathway.

For this experiment, two small groups of rats were used. One group of seven rats was trained to obtain food at *f*, and the other equal-sized group, at *b*. During the learning the number of errors (entrances) made by each animal for each kind of alley were tabulated. This type of experiment was conducted in four different years with essentially similar results. For the first three years, the maze was kept in a fixed position in the experimental room. For the experiments of the fourth year, a modification was instituted. The maze position was rotated after every third trial, through 180 deg. This was to rule out the fixed utilization of what might be called extra-maze cues, those that might be obtained from outside the maze. It was found that with this procedure, the difference in the number of *F* and *B* errors was not so great as in the studies of the first three years. It would seem from this that sensory cues from the room in which the maze was stationed formed at least a part of the basis for orientation. On the other hand, examination of the over-all results assured Dashiell that some of the orienting cues certainly lay wholly within the maze itself.

The *third experiment* involved the use of multiple-T mazes. Such a maze is pictured in Fig. 89 (right). Two multiple-T mazes were used. One was as shown in the figure; the other was an exact left-right reversal of it. These had high walls and no covering was used. Sliding doors were placed at five positions in each as indicated by dotted lines. The maze as pictured was called the *A* maze. Its mirror image was called the *B* maze. Ten rats were trained in *A*. Ten began, but only nine finished training in *B*. Both mazes were rotated 180 deg. before each fifth trial in the learning series. In the course of their training, the rats tended to make more entrances into certain blind alleys (cul-de-sacs) than others. Blind alley *C*, in each maze, was entered most frequently. Next in order were the blind alleys *E*, *A*, *B*, *D*, and *F*.

It will be noted that blind alleys *F*, *D*, and *B* all involve turns opposite the food-box direction. *C* is, of course, the one most in line with the direction of the food box (see Fig. 90).

The Main Experiments. The *first experiment* here involved the use of what was called an *open-alley* maze. It is pictured in Fig. 91 (left). In such a maze there are 20 different paths of equal length (perfectly correct) from starting point to goal or food box. An error was indicated whenever the rat took a path involving extra distance or entered into pockets (blind alleys) on any of the four sides of the maze.

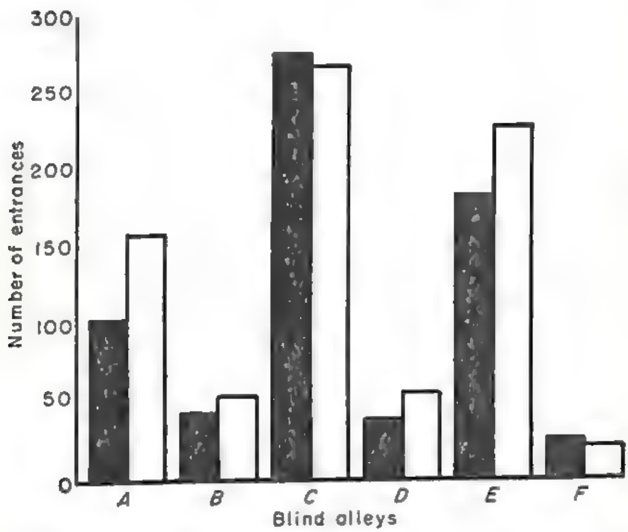


FIG. 90. Frequency polygons to indicate number of times each blind alley in the multiple-T maze was entered.

Two groups of rats, group A (96 rats) and group B (8 rats), about twelve weeks old were used. Trials were given once a day at 4 P.M. The actual course each rat took on each trial was traced on a form by the experimenter. Hunger was used as the incentive or "drive." When first a hungry rat was admitted into the maze, its behavior was of the "random," or exploratory type ultimately terminating in reaching the food box. After very few trials, the reduction in running time and distance was evidently not based upon the *fixation* of a particular pathway of runs and turns. It could only be interpreted that the animal was learning to adjust to the situation by running in the general direction of the exit to the food box. This was accomplished in one trial by one set of runs and turns, and in

others by different combinations, equally economical of distance. Thus, from the beginning, the animal was not learning a fixed pattern of runs and turns but a general orientation. Dashiell's problem lay, then, in attempting to determine more specifically what the orientation mechanisms consisted in.

The *second experiment* in this series also involved the use of the open-alley maze. In this experiment, certain sensory controls were manipulated. Visual and auditory cues from outside

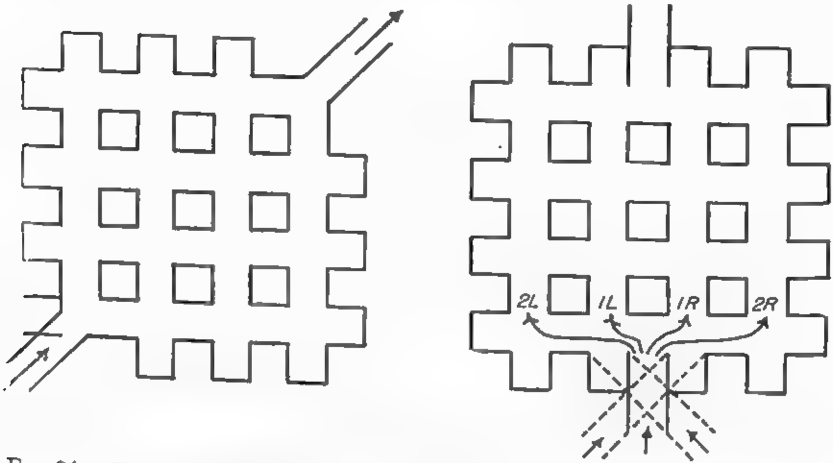


FIG. 91. Two open-alley mazes used by Dashiell. The direction of pointing of the entrance alley in the right-hand maze was changeable. (*Dashiell. Comp. Psychol. Monogr., Johns Hopkins Press.*)

the maze were checked. (1) After every fifth trial the maze was rotated 90 deg. counterclockwise. (2) After every trial the experimenter shifted his own position from right of entrance alley to left, or vice versa. To deal with visual and olfactory cues from *within* the maze, after every tenth trial the pattern of the maze was rotated, but the walls, floor, and the screen covering of the maze were retained in their original positions. To accomplish this shift, the old entrance and exit were closed and new ones opened at one or the other of the initially blank corners. This retained the identical maze pattern but shifted the orientation and the direction that the animals would have to take with reference to the established olfactory pathways on the floor.

To study the olfactory clues further, at the end of every twentieth trial the food box was thoroughly cleaned with a deodorant and replaced in position without food. For the twenty-first and forty-first trials a separate food box was put at the right-hand corner of the maze. Air from this box was forced through the maze in a general direction at right angles to the line extending from starting point to the old food box. This was done on the supposition that if food odors were directing the animal it would, in this case, be drawn over to the side from which the odor came.

Still other variations or checks were made. These related to auditory cues.

The behavior of the animals in this experiment was much the same as in the previous experiment. The errors were not increased as a result of any of the special sensory controls that were employed by the experimenter. Thus whatever the basis for orientation, it was not brought to light by any of the checks that were used. The function of orientation did not seem to be dependent upon any of the stimulus sources that were modified. Although orientation in a given situation can be said to be one that is learned, it becomes established after a few experiences and is not learned as the usual "habitual" response to exteroceptive stimuli.

Dashiell next tried to answer the question of whether orientation was based upon how the animal was introduced into the maze. To attach this problem, the design of the open-alley maze was altered to permit the deflection of the entrance alley to the right or to the left. The entrance was placed as indicated in Fig. 91 (right), and was capable of being slanted to the left or to the right as indicated by dotted lines. Ten rats were each given 50 trials in this maze with the entrance pointing toward the center of the maze. Throughout, their behavior was quite similar to that of the animals in the two previous open-alley maze experiments. Following the fiftieth trial, the angle at which the 12-in.-long entrance alley joined the maze was shifted. For the fifty-first and fifty-second runs it was angled 45 deg. to the right, so that it pointed to the left half of the maze. For the fifty-third trial it was angled 45 deg. so that it pointed to

the right half of the maze. In 29 out of 30 trials, the animals followed, for some distance through the maze, the initial orientation imposed by the angle of the entrance alley. That is, if the alley pointed toward the left, the animals began by using alleys in the left side of the maze. After choosing a left-hand or right-hand alley as the case might be, the animal followed it until reaching the end and then turned in the necessary direction to reach the food box. It became obvious to Dashiell that the pattern of the maze did not introduce enough choice situations, or thwarts, after the initial choice at the end of the starting alley had been made.

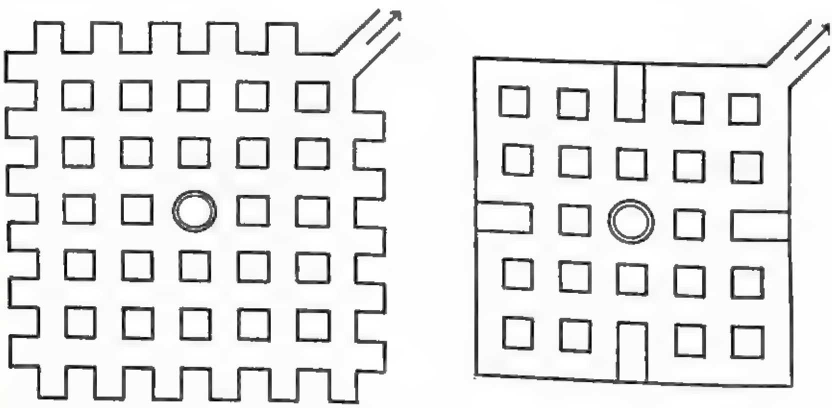


FIG. 92. Two center-entrance mazes. (*From Dashiell. Comp. Psychol. Monogr., Johns Hopkins Press.*)

The *next experiment* was designed to provide a spontaneously determined entrance into the maze. The animal had to climb up into the maze through a circular hole at the center of the maze. Two center-entrance mazes are shown in Fig. 92 (left and right).

In this experiment the maze was rotated to the right or left, assuming one of eight positions of the compass after every five trials. There also was no constancy of initial orientation of the animal as it emerged from the vertical tube and entered the maze. The question that presented itself was whether the animals tended to follow their initial head positions as they came out of the tube.

Of the northerly initial orientations, 67 per cent were followed to the opposite side of the maze where some turn had to be made. Fourteen per cent were followed halfway across. Of the easterly

orientations 73 and 15 per cent were followed all and halfway respectively. Southerly orientations were followed 69 and 19 per cent, and the westerly orientations were followed 58 and 16 per cent. The degree of similarity of the percentages for each of the four geographical directions indicates that there was no stimulus factor within the maze that had much, if any, effect in switching the animal out of the general direction in which it initially sets out. To say the least, the animal did not learn the maze as a problem capable of a single solution, a specific route to the food box.

Dashiell's final experiment consisted in excluding himself from the environment of the rats. This was a precaution instigated by the reports from other experiments, at least with such animals as dogs, trained horses, etc. In those cases, part of the animal's performance very definitely consisted in taking cues from the trainer's behavior. The original open-alley maze was used and was covered with "Synite" glass, pebbled on one side. This prevented the animal from seeing any distinct objects outside the maze. The experimenter could see the rat as it traversed the maze, owing to the nearness of the rat to the glass. Again, in this experiment, the 16 animals used demonstrated their ability to orient themselves. The animals' behavior indicated to Dashiell that their orientation was a function of the manner in which they first entered the maze.

Discussion. By use of the several mazes and procedures we have described, Dashiell obtained evidence that rats were well able to orient themselves in maze situations despite deprivation of a number of the expected sources of cues. There seemed to be no immediate way of getting at the mechanism (means) whereby orientation was accomplished. It might have been possible to bring human introspective experience to bear upon the problem, but this, if considered, was not done, owing to the possible fallacies of imputing to lower animals the mechanisms that we ourselves possess.

Dashiell turned to the literature on animal experimentation, including particularly the work that had been done in perception and learning. There was considerable variety for him to examine. The two main performances from which he sought

help were (1) homing, and (2) maze learning. As an essential part of many of the investigations in these two areas, the authors appended conclusions or suggestions in regard to how their animal subjects accomplished the behavior that they were observed to perform. Since those who worked in the area of animal psychology included not only behaviorists, reflexologists, but also mass-action theorists and configurationists, the theoretical remarks Dashiell encountered included concepts and terms such as position habits, insight, structured situations, situations-as-wholes, conditioned responses, and symbolic processes. Among the workers whose experiments were reviewed were Watson, Hunter, Pavlov, Lashley, Carr, Yoshioka, Franz, Yerkes, Gengerelli, Kohler, Helson, Higginson, Cameron, Maier, and Burtt.

From these sources, Dashiell's conclusion was that despite all the variety involved none of the experiments surveyed precisely paralleled his own. For that reason they did not supply the help needed.

Dashiell's essential conclusion, which was derived directly from his own experiments, was that the learning of a maze is not just the process of "simply integrating a chain of conditioned reflexes," not the "organizing of a simple serial pattern of discrete units." While this same conclusion might have been drawn from the experiments of the men just mentioned, Dashiell's maze experiments bring the matter to the fore and give it an emphasis that no other group of experiments have ever done. If it can be said that Dashiell's experiments have driven home the idea that learning is not a piecing together of discrete and independent elements such as movements, etc., it has made a major contribution.

QUESTIONS

1. What is meant by directional orientation in a maze?
2. What did Dashiell specifically set out to determine?
3. Describe the maze pictured in Fig. 87 and tell what learning might be expected from its use.
4. Describe the maze pictured in Fig. 89 and tell what learning might be expected from its use.
5. What is the significance of using "forward" and "backward" alleys?

6. What were Dashiell's conclusions from the findings on the maze pictured in Fig. 91?

7. Why have the animals climb into a maze through a circular hole at the center?

8. What were some of the precautions Dashiell observed in using his mazes?

9. What are some of the terms used by other investigators in connection with the behavior of animals in problem situations such as mazes?

10. What were Dashiell's final conclusions?

CHAPTER 31

REASONING IN ANIMALS

Problem solving was brought into the laboratory in animal experimentation by Thorndike in 1898. The nature of the behavior of his animal subjects led him to adopt the term "trial-and-error" as a description of the way progress in learning occurred. The term had been proposed four years earlier by Lloyd Morgan.

Since the early experiments of Thorndike, problem solving in animals has become quite common in the laboratory. A number of terms have accumulated to describe the behavior of animal subjects in such situations. Though by no means all experiments with animals have clearly required problem solving on their part, various processes or mechanisms hypothecated for such situations have been used to explain problem solution. In 1913, the animals in Hunter's delayed-reaction experiments made correct responses after delays long enough for the animals to have lost their "pointing postures." Since it was this pointing that was initially adduced to explain correct response, some other explanation had to be sought.

Among the mechanisms brought into account for animal learning were chain reflexes. Hunter, in 1920, however, stated that studies on temporal maze learning indicated that it would be all but impossible for a sequence of kinesthetic processes to explain learning in the rat. Maze learning was seen to involve cues that refer to spatial localization as well as temporal sequence. The animal must come to know, in spatial terms, where he is in a maze.

The behavior of animals seemed to involve complex perceptual processes. Animal subjects acted as though they surveyed situations before acting. It was sometimes as though reflection

and thought had occurred. Some solutions came quite suddenly at the end of a period of pause or delay. As a label for some of this behavior, the term "insight" was invoked and came to be a descriptive term quite in contrast to trial-and-error.

As in other kinds of laboratory experimentation, considerable variety in the conditions from study to study were involved. The nature of the evoked animal performance varied somewhat also from investigation to investigation. As a result, the hypotheses and descriptive terms used by experimenters displayed considerable variety. Various workers were not sure that all of them were studying the same thing, but rather felt that they were dealing with different aspects of perception and learning. On the other hand, many workers tended to conceive of all animal behavior in terms of the specific sets of conditions they themselves set up for their animals.

A criticism leveled at American animal psychologists regarding their experiments was that they did not set up situations in which their subjects were allowed to survey the over-all problematic situation in such ways that relationships might be immediately comprehended. This possibility being denied the animal subjects, the subjects had to rely upon random modes of attack, wherein irrelevancies had to be eliminated in serial fashion as trials progressed. Often it was purely by chance that the solution was arrived at.

This criticism makes us aware of the possibility that problematic situations might be classified into the two general types: (1) Those in which the opportunity to survey the over-all situation at once is given; and (2) those in which the subject is introduced to the various factors in the situation one at a time, and in a chance order.

It is needless for us to develop a complete survey of the many ways animal behavior has been studied up to now. We can proceed better if we launch into an illustrative case. For this purpose, the work of Maier has been chosen. He has studied what he calls *reasoning* in animals. By reasoning he means that a solution to a new problem is brought about by properly combining several previous "isolated" experiences.

The Problem of Reasoning. The problem for the experimenter in studying reasoning is, according to Maier, one of providing the animal subject with several isolated situations and later with a problem that involves these previous situations. Our caution here is that we must be sure to keep our sets of terms describing outward situations and internal processes separate. It is very easy to set up a number of physical situations that are included in a final situation. Such outward *combining* is concrete, and it is easy to agree whenever it happens. Whether what happens in the animal subjects from situation to situation is a process of "combining" is questionable.

There are events that happen *within* the animals placed in the initial isolated situations and in the final one. What these really are is what we should like to understand. Although it would seem that Maier is not concerned about these, we have chosen his experiments as illustrative of a mode of study. The mode provides for manipulation of physical factors and provides a way of determining how many and what kinds of situations can be utilized by the animal subjects.

Procedures. One of the simplest experiments Maier performed was the following. Three tables were used in it. The tables were connected by elevated pathways, as indicated in Fig. 93. The animal subjects (rats) were permitted (one at a time) to run from one table to the others over the elevated pathways, becoming familiar with the physical situation. This exploration period is called "Experience 1." After this the subject is placed on one of the tables, let us say *A*, and here it is fed. Maier called this "Experience 2." The animal is then put on table *B*. Presumably, if the animal has "combined" Experience 1 with Experience 2, it will immediately run from table *B* to table *A* where it was just fed. The problem can be varied from trial to trial by giving the animal its first feeding on different tables and by starting the animal from different tables.

You will immediately be able to see that the animal probably developed spatial orientation from its opportunity to traverse the whole physical setup at will (Experience 1). Whatever is meant by "combining," it is a matter of the rat still wanting

food and the experimenter's perception of the final situation in *light of* the initial experiences that involved the over-all physical setup. Actually, when the animal was exploring in Experience 1, it was developing the orientation needed to solve such a problem as Maier finally set for it. Experience 2 was obtained in light of Experience 1.

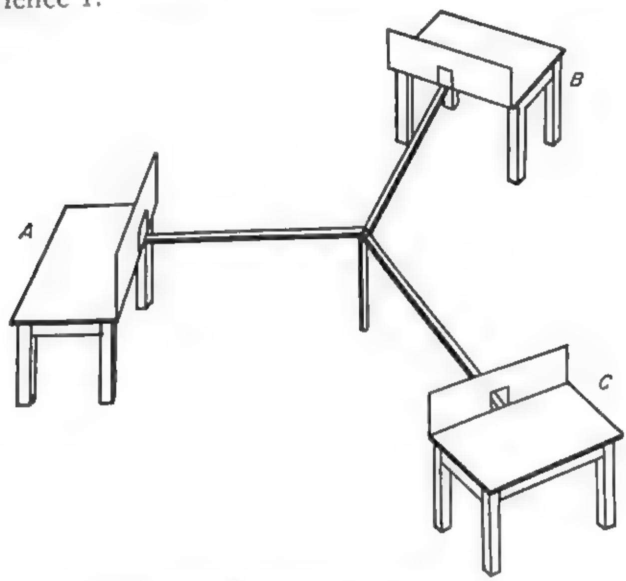


FIG. 93. A problem (reasoning) setup used by Maier.

Much more complex situations than the one just dealt with were employed by Maier. One of these is diagrammed in Fig. 94. *A* and *B* are the two halves of a large table separated by a cardboard partition. The rats are able to scramble over the partition with difficulty. An elevated maze has its origin at point *X* and its termination at point *Y*, a small wire cage on the other half of the table. Experience 1 for the rats consists in exploring the table top before the maze and starting box are connected to the table. During this "experience," the animals learn to scale the partition between the two halves of the table. Experience 2 consists in the rats being in the starting box *X* and in exploring the maze and learning to run to *Y* for food.

Finally, the rat is placed on the table at *M* with the food inside cage *Y*. At first the animal will spend considerable time and effort in attempting to get into the cage. This achievement

is not possible. Sooner or later the animal scales the partition. When this happens, the critical point is reached. The question is will the animal run directly to the box at X and enter so as to get onto the maze and be able to reach Y. Maier's rats did go almost immediately to the box when they had once scaled the partition. The rats scrambled into the box even though they had never entered it before by their own efforts. Previously

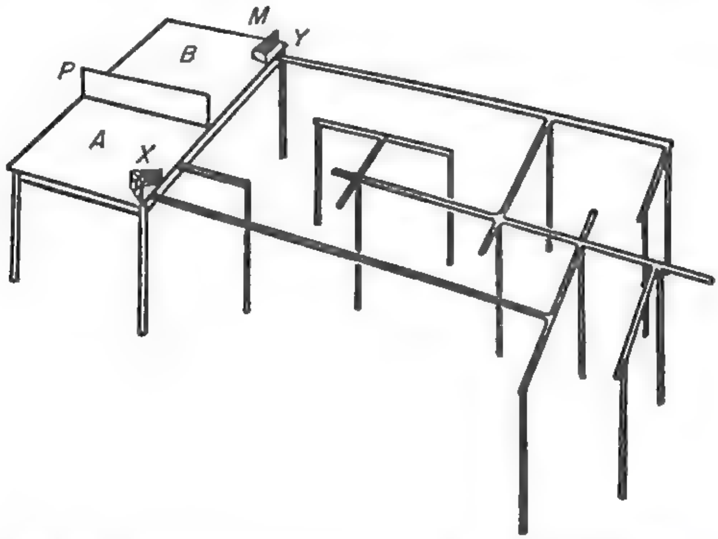


FIG. 94. An elevated-maze-reasoning setup used by Maier.

they had been manually placed in the box by the experimenter. Upon entering the box through their own efforts, they ran directly around to Y to the food. This same kind of performance has been called an example of *insight* by certain other workers.

We can say that since the experimental animals perform in such ways as to satisfy the intent of the definition that Maier set up for "reasoning," the animals displayed reasoning. We must realize that it is at once the virtue and the shortcoming of operationism that as many definitions as operations can be made.

Maier used an experimental setup in which four, rather than two, separate physical situations were provided for the animals to "combine," or "integrate" (see Fig. 95). The arrangement consists of two large tables, A and B. These are joined by pathways. A corner of table A is fenced off by a wire obstruction enclosing area F. This area can be reached only by a round-

about pathway leading from table *A* to table *B*, and thence to the enclosure. There are two pathways (exits) from each table. One, or all, of these may be blocked by the removal of sections indicated by the broken lines. For example, *DE* and *D′E′* may be removed to block passage from table *A* to table *B*. In the same way, *YZ* and *Y′Z′* may be taken out to block passage from table *B* to table *A*.

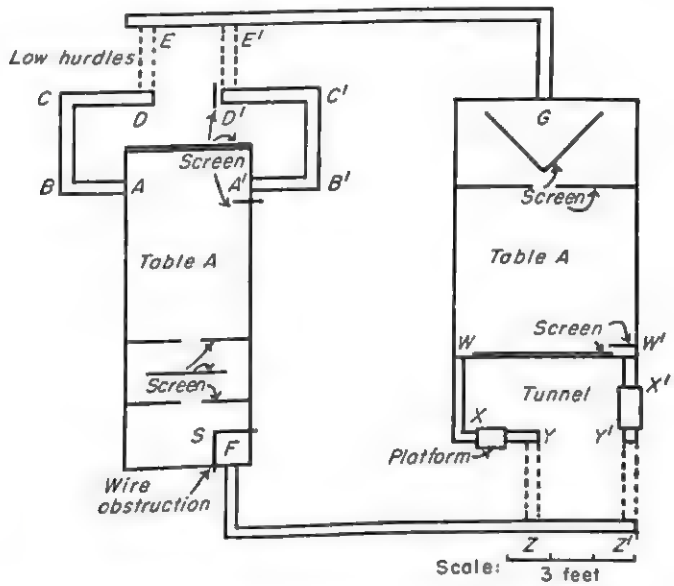


FIG. 95. A complex problem (reasoning) setup used to study normal and partially decorticated rat behavior. (*Maier. Papers, Mich. Acad. of Sci., Arts and Letters, Univ. of Mich. Press.*)

The exits have characteristic markings (landmarks) to enable the rats to orient themselves. There are screens on both tables. The screens are differently placed on the tables so as to help make it possible to visually distinguish one table from the other. The other function of the screens is to obstruct the view of the exits so that no indication of which one is blocked can be obtained until the animal reaches them.

The initial procedure consisted in allowing the animals to explore both tables, without it being possible for the animals to cross from one to the other. On each of six days, 10 min. of exploration was allowed on each table.

On the day when the rats were to be tested, four separate experiences were provided as follows:

1. A short exploration of table A, with bridges *DE* and *D'E'* removed.
2. A short exploration of table B with its exit pathways *YZ* and *Y'Z'* removed.
3. The opportunity to run from table A to point *G* of table B via the pathway with *DE* or *D'E'* in place.
4. The opportunity to run from point *Y* or *Y'* across to point *F* on table A, where food was received.

Experiences 3 and 4 were provided alternately, each three times. When the experiences had been provided as just indicated, the subject animals were individually provided with an apparent problem by being placed at starting point *S*, with the food dish at *F*, just on the other side of the barrier. To obtain the food, the rat had to be able to use the previously given isolated experiences to create a route to the food.

Two possibilities for making errors existed at the beginning. With section *DE* or *D'E'* in place, the animal had to leave table A by the path involving the proper section. Thus the *first choice points* were *A* and *A'* on table A. In the same way, exits from table B constituted a *second set of choice points*. The rat made an error if it left either table by the wrong exit, *i.e.*, one not leading to the bridge in place. A perfect run constituted the use of the right choice at the first and second set of choice points. Mere chance would account for 50 per cent of the correct selections at each choice point and for 25 per cent of the perfect runs. On this account it was necessary to try the animal several times to ascertain its ability. The combination of bridges in place was varied also from trial to trial. Each rat received a total of 20 tests. In 8 of the tests both bridges were changed from what they were the day before. In 4 of the tests only the first bridge was changed. The tests were given to 56 normal rats and to 26 rats with cortical injuries. Both the operated and the normal rats had been previously tested on a simple reasoning problem. None of the brain lesions were extensive, but were widely distributed over the cortex. The average lesion involved between

15 and 16 per cent of the surface of the brain. The lesions ranged from 8.4 to about 32 per cent, with a probable error of about 4.4 per cent.

Results. Whereas there were two types of rats used, operated and unoperated, they were classified into three groups in accordance with the types of errors made. Group A included those rats that made a higher score of correct choices at the first choice point than at the second. Group B included the animals that behaved in just the opposite way. Its score was higher for the second choice point. Group C included those animals that made

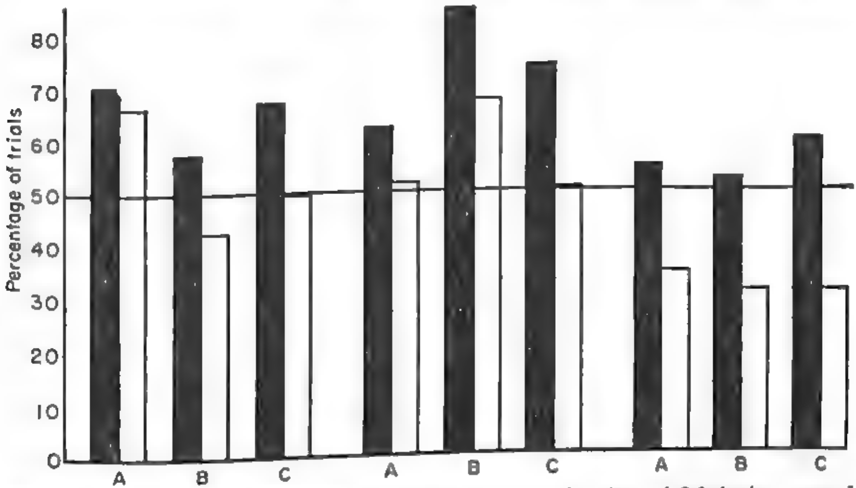


FIG. 96. Frequency polygons to indicate the behavior of Maier's normal and partially decorticated rats. Black rectangles represent normal rats; white, the operated rats. (See text for further details.)

scores within one point of being the same at both choice points. Figure 96 indicates the results obtained with the operated and normal rats in each of the three groups.

Maier made a number of other comparisons between the normal and operated animals, but the results are too lengthy to report here in tabular form. We can only enumerate *some* of the more general findings.

1. Normal animals are able to "combine" the four separate experiences, that is, make percentages of perfect runs in the test situations that are decidedly above chance.

2. The scores for the operated rats are very little above chance, and it was concluded that the operated rats failed to solve the

problem as a whole. They actually were not making purely chance records, but were reacting to parts of the problem situation. This, of course, was inadequate so far as perfect runs were concerned.

3. Normal rats in groups A and C scored better than chance at the first choice points for all types of changes of the bridges. Normal rats in group B scored slightly below chance when the first bridge was changed, and when both bridges were changed.

4. Operated rats made perfect scores in excess of chance only when the second bridge was changed. As a group, these rats solved the problem mainly by adjusting to a change at the second point. In the other instances they repeated the performance of the previous day or else made a chance score. Such an adjustment to part of the general situation resulted in perfect runs above chance.

5. With their reduction in ability to integrate, the operated rats tended more and more to utilize learned reactions rather than purely random behavior. The most salient learned response stemmed from experience of a previous route to food.

QUESTIONS

1. Who was the individual usually credited with doing the first laboratory experiments with animals?
2. What is meant by "trial and error" learning?
3. What other kind or kinds of behavior beside trial-and-error performance have been observed in animals?
4. State a criticism that has been leveled at American animal psychology.
5. How does Maier define reasoning?
6. Does Maier's definition of reasoning describe what is going on in the animal or what is happening outside (in its overt behavior)?
7. What is the essential procedure by which Maier tests reasoning?
8. Do animals indicate that they reason?
9. What difference was found between operated and unoperated rats in Maier's experiments?
10. Were chance successes taken into account in Maier's experiments?

CHAPTER 32

TOKEN REWARDS IN ANIMAL LEARNING

What are loosely known as motivational factors play a leading part in the determination in the behavior of the organism. Whether an animal is hungry and whether some specific performance is followed by food or not are of crucial significance in the subsequent behavior of the animal in similar conditions. Food as an incentive has been taken as axiomatic. Animal experiments in learning have therefore utilized various kinds and amounts of deprivation (food in particular) and, on the other hand, various amounts of incentive or reward. One kind of analysis into the problem of reward is represented by a group of investigations in which the efficiency of indirect rewards has been studied.

A number of years ago Williams showed that a rat will improve in maze running when the immediate goal is the food-associated incentive of a previously learned, visual-discrimination habit. McCulloch reported that an infant chimpanzee learned a visual-discrimination problem where the incentive was a paper towel previously neutral or negative, but which had come to be associated with the highly desired activity of running to the experimenter's arms. Wolfe, in 1936, used small disks (tokens) to which chimpanzees were initially quite indifferent, and by means of associated training, these became the incentive for performing work. He not only showed that the tokens served to elicit single work responses, but to elicit a succession of such responses before the tokens were exchanged for food. Other workers, Nissen and Crawford, have also used food tokens to evoke certain kinds of social behavior in chimpanzees. Ellson trained dogs to exchange rubber balls for food. Subsequently the balls became the incentive for performing a simple work response, the depressing of a lever to release balls to the animals.

Most of these studies involve the performance of an already learned response. Few experiments have involved the development of new habits as the observed function. The investigation that we are going to deal with in this chapter is that of Cowles. He believed that more experimental information on the problem of substitute rewards in learning was needed. The significance of indirect or vicarious rewards may become more fully evident when such rewards lead to the formation of new habits than when they merely prompt repetitions of a previously learned performance. The first part of Cowles' investigation was concerned with demonstrating whether the animal will learn to choose between two or more alternatives that are almost equally preferable at the beginning of training, when the incentives are food tokens. The second part of the investigation was an attempt to determine quantitatively the approximate value of the token reward as a reinforcing factor in the acquiring of habits. In addition, he wanted to determine any possible qualitative differences in the activity manifested toward token rewards and other incentives.

The various general conditions used in the investigation involved comparisons between food-token rewards and food itself, for example, (1) the receiving of food for a correct response not requiring the exchange of a token for it; (2) food tokens immediately exchangeable for food after each correct response; and (3) "reward" with tokens which had no exchange value and which involved special training to teach the animal that fact.

Cowles used five young chimpanzees: three females, Alpha, Bimba, and Bula; two males, Frank and Moos. The three females were 5 years 4 months, 7 years, and 5 years 8 months old, respectively. The males were 6 and 9 years old, respectively. All the animals had known the experimenter for two or three years, and all except Moos had previously served as subjects for him.

Apparatus and Procedures. The apparatus was housed in two adjacent rooms, as indicated in Fig. 97. In one room was a large circular cage. At one of two grills, in the cage wall, 90 deg. apart, was placed a vendor for tokens that either automatically or by the experimenter's switch control delivered food (a raisin)

to the animal whenever a food token was placed in a certain slot. A shutter was placed in front of the vendor so that the animal could be denied access to it in accord with the plan of the investigation. At the second grill, a work apparatus was placed when needed. A large opaque vertical screen concealed the experimenter. Two work devices were used at various times. One was for weight lifting; the other was a sliding tray in a grooved runway used for "pulling-in" responses. Attached to the tray was a light cord by which the animal was enabled to

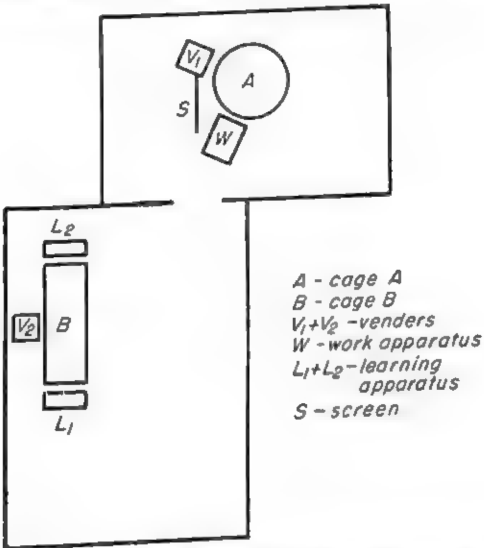


FIG. 97. Floor plan of the two rooms with their cages, food vendors, and problem apparatuses used by Cowles to study chimpanzees' responses to food tokens.

draw the tray to itself. In the second room was a large rectangular cage having a coarse vertical grill of ½-in. steel bars at either end. The animals could reach through this grill and readily manipulate apparatus outside the cage. A wooden barrier was placed in front of one grill or the other so that only one grill was open at a time. In front of each grill was a fixed framework supporting a learning-problem device near enough for convenient manipulation by the animal. Each learning device carried a drop door to close the device from view between trials. A screen was also used to shield the experimenter from the animal's view at each of the learning devices. A small grill

was located on one side of the cage to permit access to a food vendor quite similar to the one just described. This vendor could be excluded from the animal by a drop door when desired.

The only food that was used in the investigation was large, dried Malaga raisins containing seeds. The animals always seemed eager for them throughout the experimentation that lasted a number of months. To begin with, the raisins were cautiously received but at no time were they ever refused. The animals quickly came to like this new food. The tokens were small, unbreakable disks 1¼ in. in diameter and 1/12 in. thick. These were small enough to be carried in the animal's hand from one room to the other and yet thick enough to be readily picked up from the floor by the animals or separated from one another whenever they adhered together.

All the animals, except Frank, had had some experience with tokens used to exchange for food in a vendor. Frank learned to use the tokens in one short demonstration period. After a few trials in testing the willingness of the other animals to exchange the tokens handed to them for raisins, the use of the work device was begun. When the animal demonstrated that it would willingly work for a food token immediately exchangeable for food, the necessity of collecting two or more tokens was then required before exchange could occur. To prevent immediate exchange, the experimenter closed the shutter in front of the vendor. Gradually the number of tokens was increased depending upon the willingness of the animal to perform. Verbal encouragement was often used to return the animal to work after unusually prolonged interruptions brought about by play, outside distractions, etc. The objective was to arrive at a practical limit beyond the number of tokens to be received later in another learning situation. The maximum number of tokens to be worked for before exchange could occur was arbitrarily determined by the animal itself. Repeated pauses longer than half a minute, even when verbal encouragement was used, set the limit. When the animal came to work quite readily and consistently for such a number of tokens, it was given opportunity to acquaint itself with the second room and second cage (B). When in this cage, for example, it was given a few tokens that could

later be used in the vendor of cage A. Now the animal was ready to begin regular trials in cage B with learning apparatus.

The general procedure for learning under three different conditions (B, C, and D) was as follows. On each learning occasion the animal was taken to cage B and given a number of trials without pause. Generally 20 trials were used. For each response to the proper stimulus, the animal received a token. After the series was finished, the animal was led to the next room and allowed to exchange the tokens that he had earned for raisins in the cage-A vendor. The time elapsing was about 50 sec., and the animal had to walk a distance of about 35 ft. The animals varied in the number of tokens they earned, and if the series of 20 uninterrupted trials was at first too great, it was reduced to 10. The animal returned to cage B for the second set of 10 trials after receiving its raisins in cage A. After this series it was allowed to exchange its tokens and then was taken to its living quarters.

We shall be able to describe only a few of the learning experiments. After a beginning acquaintance with the fact of tokens being exchangeable for food (procedure A), the first experiment that Cowles used involved what was called a two-position learning apparatus (procedure B). It consisted of two aluminum cups anchored each within its own wooden box and covered by a lid hinged at the rear. To gain access to it, the animal had only to lift the lid. In this two-position problem, one or the other of the cups (let us say the right-hand one) was baited with a food token in each trial until the animal chose this cup consistently. The other cup was then baited instead until the animal shifted its previously learned choice and now chose the second cup each time.

After this, the animal received 12 preliminary trials, each container being baited a number of times, but in an irregular order. In case of error the animal was allowed a second choice. For what were called the regular learning trials, the animal was given either one or two 20-trial periods per session and was allowed to exchange its tokens at the end of either 10 or 20 trials. These differences were made in adjustment to the apparent ability of the animals to learn or to be motivated by token rewards. At

each experimental session a reversal in the cup that was baited was made. The sessions were from 6 to 18 hrs. apart.

In his second learning experiment, Cowles used a five-position learning apparatus and interspersed food with token rewards (C). In this experiment a single cup was baited in each trial. A series of trials were run, and each time the animal chose the right cup (baited cup), a food token was given. This was exchangeable for food only at the end of a 20-trial series. Another series was run in which the animal was rewarded by food in each trial.

One method of baiting the cups (called the "criterion" technique) was that of using the same cup each time until the animal consistently chose that cup. Following this sort of behavior, another cup was chosen and baited each time until the same consistent behavior was elicited. This was done for each of the five cups. In this series, either a raisin or a token was the reward. In all, five different place habits were formed and broken for both the token and the raisin.

The other method, called the "equal periods" technique, involved baiting the five cups in random order for a series of 20 trials. Two 20-trial periods were used for each reward, the token or the food itself. These sessions were balanced between mornings and afternoons. Each cup was thus baited in 80 trials. For each incentive, there were two 20-trial periods, one in the morning and one in the afternoon. For Bimba, food tokens only were used.

In a third experiment, Cowles used a visual discrimination problem and gave food tokens only (D). In later problems (E, F, G) two kinds of tokens were used: one that could be exchanged for food, and the other that could not. In these situations it was typical for the nonfood tokens as well as the food tokens to be removed from the learning apparatus. The animal generally left the nonfood tokens in cage B when it went to cage A to exchange the food tokens for raisins. If any nonfood tokens were carried to cage A, they either were used after the food tokens, or not at all. If a nonfood token was placed in the vendor, the experimenter was able to prevent the vendor from delivering food for it.

Results. Although the actual data constituting the findings are too detailed to be presented in the space we have, there were several results that may be summarized.

1. Cowles verified the earlier findings that chimpanzees will work for tokens exchangeable for food. He found that they would not only work for single tokens bringing immediate reward, but would (after certain training) quite willingly work for as many as 10 to 30 tokens earned one at a time, exchangeable ultimately as a group. Definite individual differences were manifested in this respect. Cowles attributed these to what he called emotional factors.

2. Cowles showed that tokens exchangeable later for food were an adequate incentive for learning and "remembering" ("retaining the habit") several different problems; namely, (1) simple left-or-right position habits; (2) five-position habits; (3) visual size discrimination; (4) discrimination of color patterns; and (5) delayed responses. All of the animals in the first four problems met the prescribed criteria for mastery that involved behavior significantly beyond chance expectations.

The learning of these habits showed definite progress during single experimental sessions, prior to the receiving of the reward of food from exchange of the food tokens. Food tokens functioned as rewards for intervals as long as 2 min.

All of the five problems that Cowles used were mastered by at least one chimpanzee who had not previously secured food either in the learning apparatus or in the experimental quarters. This means that each problem was mastered by a subject for which the only object associated directly with food reward was the tokens themselves. Actually "full-fledged" habits were learned (even though they required a number of trials) solely upon the basis of food tokens as differential reinforcing agents, prior to the receiving of food itself as the reward.

3. Cowles also found that the animals could learn to respond differentially (choose one, reject the other) to two kinds of tokens, one kind exchangeable for food, the other with no exchange value. The process of learning this sort of discrimination came first by way of immediate exchange for food in the case of the one but not the other. Gradually grouped, exchange of tokens

earned one at a time could be comprehended by the animals for the food token and not expected of the other type.

4. The reward value of the food token was compared with the results of the nonfood token. For example, in a number of trials "to criterion," when the alternate stimulus was "rewarded" with nonfood tokens, the correct response was learned neither more nor less rapidly than when food tokens alone were used. The rate of learning did not show the superiority of either, but the accuracy of responses taken as a total showed that rewarding by food tokens led to a greater per cent choice of the food-token-baited stimulus than when the other method was used. Even though the animals continued not to try to exchange the nonfood tokens for food throughout a learning period, their scores on nonfood-token trials were decidedly above chance in number.

5. In some problems and under some conditions it was found that reward with food directly was more effective for learning than reward with food tokens. This was true in the first problem when the animals were just learning the value of food tokens. In the second problem, receipt of food directly was only slightly superior to receiving food tokens. General motivation seemed to be higher during learning where food and food tokens were alternated, as in contrast to the trials when food tokens and nonfood tokens were alternated. It must be remembered, however, that in the food to food-token comparison twice as much food was received.

6. In some tests, all three "rewards" were used, and a hierarchy of effectiveness was demonstrated. Naturally food was more effective than food tokens. Food tokens were more effective than nonfood tokens. The latter were not entirely ineffective, although far less effective than food tokens.

Cowles pointed out that there are several ways of regarding a token which through use has been connected with a primary reward such as food. It may be looked upon as (1) a *sign* or symbol of something else going to occur as, for instance, the obtaining of food or some activity privilege; (2) a tool—a device, a *means to an end;* (3) something that on account of past experiences has now developed value of its own. It was pointed out

that a choice between the alternatives could not be made upon the basis of the experiments carried out. Cowles did, however, discuss at length the three kinds of properties and cite instances of animal behavior which point in the three directions respectively.

QUESTIONS

1. What is a token reward as indicated in the present chapter?
2. Cite two animal studies in which indirect rewards were found effective.
3. In what respect did Cowles intend his investigation to differ from previous studies with token rewards?
4. What three kinds of rewards were used in Cowles' investigation?
5. By what means were the chimpanzees able to exchange tokens for food?
6. What is meant by Cowles' "criterion" technique?
7. What was the "equal periods" technique?
8. How did the animals treat the nonfood tokens?
9. Did each token have to be exchanged for food as soon as received? If not, what limitations in time or accumulation were found?
10. Briefly summarize Cowles' findings.

CHAPTER 33

CONFIGURATIONAL LEARNING IN THE GOLDFISH

In animals below the level of mammals and birds, matters of problem solving and preferences for various stimulus conditions have not been very extensively studied by psychologists. This is particularly true of experiments to test the existence or absence of ability to respond to specific elements in complex situations. Early in the history of the study of discriminative response and of learning, investigators overlooked the possibility that context had a great deal to do with the way certain specific stimulus objects were responded to. With the advent of a school (Gestalt psychology, configurationism) that sought to revamp psychological thought in regard to wholes and parts, some attention to context in animal experimentation began to be given. At present, recognition of context in one way or another is so commonplace that it is difficult for a beginning student to realize how remote such a consideration used to be. This is not to say, however, that all or even much experimentation follows in a consistent or full-fledged manner the principles of the configurational school.

In the investigation about to be described, we have a good example of a case in which the problem of the existence of configurational response was studied in a species of fish. The study is a good example of the kind of work prompted by configurational and organismic viewpoints. Furthermore it provides an understanding of how *aquatic* species may be handled in a laboratory-problem situation.

The study in question is Perkins' investigation of certain problems evolving from his work with Wheeler on learning in the goldfish. In that study, the authors trained goldfish to choose a light in relation to two others of differing intensity. For

example, when a fish was trained to choose the mediumly lighted compartment in a group of three, it was able to choose correctly when the absolute levels of all compartments were raised or lowered, or when the relative illuminations were shifted.

The study left certain problems to be investigated. In the work to be described in this chapter, Perkins set out to solve some of these problems. They were (1) whether the fish were responding to intensity alone or also to wavelength; therefore, Perkins, in the present study, eliminated wavelength changes that accompanied shifts in illumination to find out whether the fish would respond as before; (2) whether the fish preferred compartments of medium illumination rather than those of lower or higher illumination. It was also desirable to study (3) behavior of the fish in constantly changing sets of absolute-intensity relationships between the three compartments, as well as constantly shifting positions of the three.

Apparatus and Procedure. For this experiment a three-choice arrangement was provided. It consisted of an arrangement such

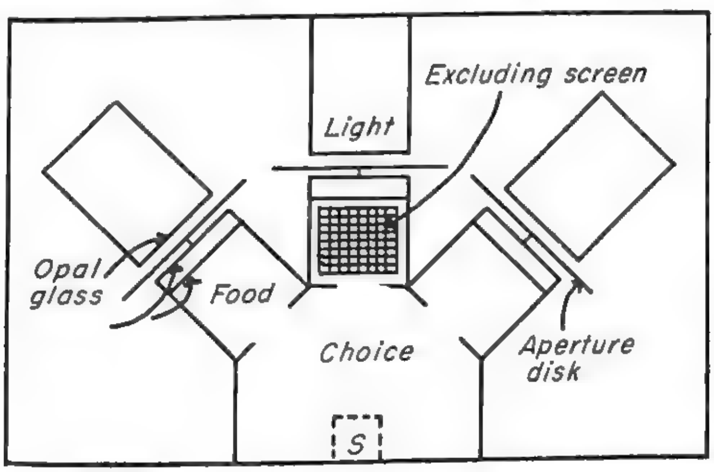


Fig. 98. Arrangement of compartments for goldfish-learning experiments. *S* is the starting cage. (*Perkins. By permission of J. exp. Psychol. and Amer. Psychol. Assn.*)

as pictured in Fig. 98. In the diagram is a choice compartment, from one side of which led the three symmetrically grouped food compartments. Their axes converged in the choice compartment, so that from a mid-position in it the three doorways were

easily visible and about equidistant. Each food compartment was illuminated by an opal-glass wall at the end opposite the entrance doorway. The lamps illuminating each of the food compartments were confined in a box (lamp house) so as to prevent distribution of stray light to unwanted areas. The intensity of the light illuminating the opal-glass wall of the food compartment was controlled by selection of an aperture in a revolvable disk. Light from the lamp house was diffused before passing through the particular apertures in use and diffused again by a second plate of opal glass before reaching the opal-glass wall of the food compartment. Any one of five apertures could be chosen by merely revolving the disk. The absolute brightness of the opal-glass wall food compartment could, in this way, be varied from 10.55 c./ft.^{2} (candles per square foot) (aperture 1) to 0.020 c./ft.^{2} (aperture 5) with the following intervening steps, 4.8, 0.73, 0.21 c./ft.^{2} In addition, the disk could be turned to involve no aperture (zero aperture). Some stray light, of course, was involved when the zero aperture was used, for it could not be said that the food compartment was in absolute darkness.

To exclude the fish from the food compartment when desired, a wire-mesh screen was hinged to the floor of each food compartment at the edge near the entrance. A black thread attached to the opposite side of the screen was used to lift it (swing it upward). This forced the fish out of the compartment without hurting it and without unduly disturbing it.

The entire arrangement was submerged in a large aquarium to a depth of 5 in. The aquarium was situated in a darkened room. A cubical wire-mesh cage 6 by 6 cm., with no top or bottom, rested on the floor of the choice compartment opposite the entrances to the food boxes and symmetrically positioned with reference to them.

At the beginning of each trial a fish was placed in this cage. Its release was effected by merely raising the screen, whereupon the fish was free to swim about in the choice compartment.

At the opposite end of the large aquarium and independent of the arrangement just described was a group of individual wire cages in which the animals lived during the investigation. Dippers were used to lift the fish from their residential cages to the

starting cage in the choice compartment. All the fish were nearly the same size and, as nearly as could be ascertained, from the same school. The fish were never fed in the residential cages, but always in the food compartments.

Equal amounts of commercial fish food were provided in each of the three compartments for each trial. If the fish entered the "correct" compartment, it was left there about 20 sec. to eat. If at the end of this time the fish had not swum back out into the choice compartment, the experimenter lightly tapped the top of the food compartment. This aroused the fish and permitted it an opportunity to leave the food compartment by itself. When the fish entered an "incorrect" food compartment, the experimenter immediately pulled the thread attached to the wire-mesh screen already described and caused the fish quickly to leave the compartment. The experimenter then lowered the wire starting cage over the fish and guided it to starting position for a new trial. Each fish was placed in the apparatus for 10 min. each day or until it had made five responses somewhere within that time. A response was recorded when a fish entered a food compartment, at least as far as the rear fins. The problem was considered learned when the fish made 10 correct responses out of 10 consecutive trials. Forty-eight fish were used, and they made, in all, 3,812 responses.

The fish were divided into four groups which Perkins called (1) the dynamic group, (2) the static group, (3) the dynamic control group, and (4) the static control group. By the dynamic group was meant the group for which a constantly changing relationship between the intensities and positions of the low-, medium-, and high-intensity food compartments was used. After each response of a fish in this group, a change was made not only in the relative positions of the three (low, medium, and high) compartments, but in their absolute intensities.

The arrangements for the fish in the dynamic group were as shown in the columns of Table 17, in which *M* is medium, *B* is bright, and *D* is dim.

No pattern of intensities occurred more than once in any day's trials. And on each day a different order of patterns was employed by starting from a different place in the series. Nine fish were

used in the dynamic group. Three of them (in subgroup A) were trained to secure their food from the bright compartment. Subgroup B, containing another three fish, was trained to obtain food from the medium compartment, and subgroup C was trained to obtain food from the dim compartment.

TABLE 17

Combination	Left	Center	Right
I	*M* (3)	*B* (1)	*D* (5)
II	*D* (4)	*M* (2)	*B* (1)
III	*D* (5)	*B* (1)	*M* (4)
IV	*B* (2)	*D* (5)	*M* (3)
V	*M* (3)	*B* (1)	*D* (4)
VI	*M* (4)	*D* (5)	*B* (2)
VII	*B* (1)	*M* (3)	*D* (2)
VIII	*M* (3)	*B* (2)	*D* (4)
IX	*D* (5)	*M* (4)	*B* (3)
X	*M* (2)	*B* (1)	*D* (5)

The fish in the static group were also divided into three subgroups. Fish in subgroup D were trained to choose the bright compartment (lighted by aperture 1) no matter in which of the three positions it might occur. After this problem was learned, a test was made with a combination of compartments much reduced in intensity, namely, lighted by apertures 4, 5, and 0, in which 4 would be the bright compartment. When the fish had made correct choices in each of 10 consecutive trials, another combination of intensities was introduced, namely, those provided by apertures 2, 3, and 4. After the necessary perfect performances, the subjects were given a rigid test in which 10 combinations with novel relationships were used.

Subgroup E constituted the fish that were trained to choose the mediumly lighted compartment (aperture 3). When the problem was learned, the fish were given the new combination, 4, 5, 0, and training was discontinued. As a test, after mastering this combination, the intensity pattern, 1, 2, 3, was given these fish. The test given the subgroup B of the dynamic group was then given.

Subgroup F was trained to choose the dim compartment (lighted by aperture 5). After learning this, the fish were given combination 4, 5, 0, in which 5 was the medium compartment. Then combination 1, 2, 3 was given. But no final test, such as for subgroup D and subgroup E, was given.

Thirty fish were used in the control group that was separated into two main divisions. These were divided into six subgroups. In this part of the experiment, the object was to determine the preference of fish for the compartments of different illumination levels. Hence there was no training series and no "punishment." Each fish made 20 responses with medium, bright, and dim compartments, which were changed in position after every response.

The one main group was the dynamic control group. It was composed of five fish. These five fish were given the combinations used in the dynamic experiment. That is to say, the combinations differed in both absolute intensity as well as relative position after each response. The static control group was composed of 25 fish. Each fish in this group was presented with only three intensities of compartment. This main group was divided into five subgroups. The subjects in subgroup G were given combinations in which apertures 1, 3, and 5 were used. Subgroup H was given combination 2, 4, 0. Subgroup I was given 1, 2, 3, and subgroup J was given 2, 3, 4. Subgroup K was given 3, 4, and 5. This is made clear by Table 18.

TABLE 18

TABLE 18. ILLUMINATION COMBINATIONS FOR STATIC CONTROL GROUP

G	1		3		5	
H		2		4		0
I	1	2	3			
J		2	3	4		
K			3	4	5	

Quantitative Results. Perkins included in the report of his investigation a number of graphs indicating the actual behavior of individual fish as they progressed through their learning series. In some of the graphs (learning curves) there were plateaus followed by very definite sudden spurts in improvement, which was maintained from then on. This sort of behavior is the type taken to signify "insight" by configurationists.

Perkins also included in his report diagrams of the paths traversed by individual fish. These were valuable in demonstrating the reconnoitering done by the fish before making a final choice. The paths do not, of course, indicate the speed of movement involved, but we get that from the verbal descriptions of behavior that Perkins provided. In them, it was typical to read that after reconnoitering and surveying each of the three choice possibilities, the fish would make a dart for the proper compartment. The reconnoitering was, of course, taken as one kind of evidence to indicate that the choice was formed on the basis of the total situation rather than with reference to the absolute brightness of a single compartment.

Qualitative Results. None of the fish responded to a compartment as of a fixed intensity level. The fish swam around the choice compartment before making a selection. In most cases it not only circled the choice compartment, but nosed into the three doorways. This gave indication that the fish was not responding to one light level alone, but to a given level in relation to the others. The subjects often nosed into the compartment next in level to the correct one as though making a comparison of two levels of light before making a final choice. In many cases, the fish, following a complete circuit of the choice compartment, made a sudden dart into the correct compartment. This kind of behavior was most striking when it occurred in a new situation in which a transposition[1] of intensities was presented.

In both the present investigation and in the earlier one of Perkins and Wheeler, it was characteristic for a fish that went into the wrong compartment (made a wrong choice) to dart out of it and into the right one when the wire screen was lifted to

[1] The term "transposition" is used in two ways in such experiments as this. To transpose means to retain the same pattern but to shift absolute values. In music, for example, a melody can be retained even when the pitch is raised or lowered for all the notes. The second usage of the word transposition refers to the behavior of the subject (animal or human) in which the response is what it should be were the subject to perceive the similarity of pattern despite the shift in absolute values, etc. In such cases, the vernacular has it that the subject "transposes."

force it out of the compartment. In only a very few instances did a fish dart into another wrong compartment when forced out of the first wrong one. This performance naturally occurred too rapidly to allow the experimenter to do anything about it, even if he wished to do so. Such responses were not included in the treatment of the data. Notwithstanding their not being counted, the behavior was suggestive of the emergence of "insight" in the fish and deserves mention here in describing the qualitative features of response.

Another type of response was often observed early in the investigation. It was the backing into and the coming out head first from a given compartment. This was particularly true of the fish that were to choose a brightly lighted compartment. It thus appeared as though the fish were acting against a "fear." Tensions both toward and away from the goal seemed to be present.

Conclusions. Detailed specific conclusions are not so important for our purposes as the thinking and techniques involved in experiment construction. Nevertheless, it is in order to list at least some of Perkins' conclusions. Some of the conclusions were as follows:

1. Goldfish readily learn to discriminate between different levels of light intensity incidental to the problem of selecting compartments containing food.

2. The fish were able to discern a "constant relationship" between the illumination levels of the three food compartments. This held true when the successive experimental situations were shifted up or down the intensity scale after each trial in a manner avoiding any repetition of conditions on any single day.

3. Some evidence was provided which indicated that the brightness steps at the higher levels were harder to discriminate between than those lower in the scale.

4. There was evidence that the fish tended to choose the compartment of middle intensity, but if the intensities in the experimental combination were widely divergent, a preference for the dim compartment developed.

5. Perkins interpreted the behavior of the fish in configurational terms rather than in terms of trial-and-error.

6. Perkins also saw in the behavior of the fish the exemplification of the "law of least action." [1]

QUESTIONS

1. What is meant by *configurational* learning?
2. Is Perkins' investigation mainly an experiment in learning or in perception?
3. Beside being a configurational experiment, what other kind of a demonstration was it pointed out by Perkins to be?
4. Describe the apparatus used.
5. In what way or ways was the present investigation an improvement upon the previous study by Wheeler and Perkins?
6. Into what groups were the subjects divided, and how was each group studied?
7. Describe Perkins' actual experimental procedure with a subject.
8. Upon what relational principles did Perkins manipulate the light intensities in the food compartments?
9. What behavior was taken to be evidence of the development of "insight"?
10. What were some of the evidences for believing the goldfishes' response to be configurational?

[1] You, as a student of experimental psychology, will do well to look into the concepts of least action, maximum work, etc., formulated by Wheeler. They are given in his textbook, *The Science of Psychology*.

PART VII

LEARNING AND MEMORY

CHAPTER 34

THE VOLUNTARY CONTROL OF A REFLEX

A number of phenomena that represent change in the organism are taken to be examples of *learning*. In contradistinction to these are many other kinds of change classified into various categories such as fatigue, impairment, disorganization, deterioration, acclimatization, tolerance, and aging. Needless to say not all categories of change are mutually exclusive, nor is it possible to distinguish some of the phenomena in them from learning. In general, learning is the change which involves some recognized goal, some end desired either by the changing individual himself or by the one imposing certain environmental conditions upon him. Some learning is merely incidental to reaching a goal that apparently has nothing to do with the changes taking place. Such learning may be spoken of as unintentional. An individual's perceptions of external situations change from time to time, and it is difficult to know whether to deal with such changes from the standpoint of perceptual organization or from the standpoint of learning itself.

Several kinds of problems stand out in learning. The following classification may be taken as an example, although it is not to be looked upon as a rigid and formal affair:

1. The learning that arises incidental to repeated stimulation, where no obvious volition to learn is involved, and in which the attention may or may not be directed on the learning "problem." This is called conditioning and occurs best when the same, or similar, external conditions recur. Conditioning may suddenly take place or may require many repetitions. Only a few occasions, or possibly a single occasion, may be necessary to establish the new behavior. This is particularly true in the case of unusual

conditions that upon examination may be found to involve threat, etc.

2. Acquisition of manual or other motor skills. These often require extreme precision, great speed, or a rigid relation of timing and movement to intended end-results.

3. Rote memorization. In this category, we have the development of response connections with material that previously had no such connections. This learning differs to some extent from conditioning, particularly inasmuch as the connections are intended by the subject.

4. Acquisition of voluntary control of reflex responses. In this, the process of conditioning is involved. Such learning may involve some aspects of the factors of conditioning and some of those of rote learning of material.

5. The development of understanding (slow) and the emergence of insight (sudden).

One of the crucial aspects in learning concerns the manner of presentation of the stimulus situation. It may be presented all at one time or it may be presented piecemeal, as when a subject traverses a maze. In the latter there is no opportunity, at first, to grasp the whole situation. On this account, behavior can be little better than trial-and-error. On the other hand, even in such situations, learners do utilize "hypotheses." Meager as the cues may be in the early stages, cues are taken and utilized in line with certain suppositions, tendencies, etc., possessed by the learning organism. Behavior is seldom, if ever, fully chance in character.

The problem that is to occupy our consideration in this chapter concerns the possibility of gaining voluntary control of functions not ordinarily initiated and regulated by volition. Such functions are spoken of as reflexes. Put in simple terms, the question is whether we can learn to control reflexes. It is important to establish whether or not this can be done and how readily and well it may occur. If the establishment of voluntary control of systemic processes can be shown, then it can be expected that the individual possesses the ability to correct a number of unfavorable reactions developed unintentionally during the course of response to external conditions. A great many faulty organ-

ismic reactions are simply learned responses. Whereas conditions can arise to induce unquestionably unfortunate systemic reactions, the correction of these faulty reactions has never been nearly so sure, if at all possible. Faulty behavior represents potentially a variety, but desirable reactions are often represented in only one kind of performance for the given stimulus situation.

The specific investigation for our consideration is that of Hudgins in which he attempted to condition the pupillary response of the eye to verbal stimuli and to other forms of stimuli arising from voluntary movements of the individual. Cason had earlier demonstrated that pupillary responses could be conditioned to the sound of a bell. Hudgins wanted to ascertain whether effective stimuli could come from the individual himself and thus constitute a voluntary means of self-conditioning. We shall postpone, to the end of the chapter, a consideration of what is meant by the term "voluntary."

Apparatus. The apparatus consisted essentially in an enclosed light source, an optical system directing the light into the eye of the subject, and a means of indicating pupil diameter, most of which was enclosed in a cabinet, as shown in Fig. 99.

For the light source, a 100-watt incandescent lamp aided by a reflector was set up at an optical distance of 136 cm. from the eye. The light left the lamp house (bottom of the cabinet) through a lens some distance above it, which focused it onto a 90-deg. prism. From the prism the light was directed horizontally into the subject's eye. The intensity of the light at the eye was controlled by an iris diaphragm between the lamp and the lens. By means of a lever control, the experimenter could adjust the size of the iris diaphragm from a mere "pinpoint" to a maximum of 36 mm.

Only two intensities of light were actually used in the experiments. The high intensity was provided by a diaphragm aperture of 28 mm. and the low intensity, by a 4-mm. opening. The two intensities were 37.7 and 3.5 meter candles at the eye. The high intensity was used in the training procedures, and the low intensity was used in the tests.

The target (disk of light) which the subject saw as he looked toward the prism was supplied with a set of "cross hairs." These

were formed by two fine wires at right angles forming a cross dividing the disk into four equal parts. The subject fixated the intersection of the cross.

Determinations of pupil sizes were made with a Bausch and Lomb 5-power telescope having a micrometer ocular. The telescope was immediately in front of the subject and slightly above his line of regard. One of the hairs of the micrometer scale was fixed, and the other one was movable. Not only the fixed

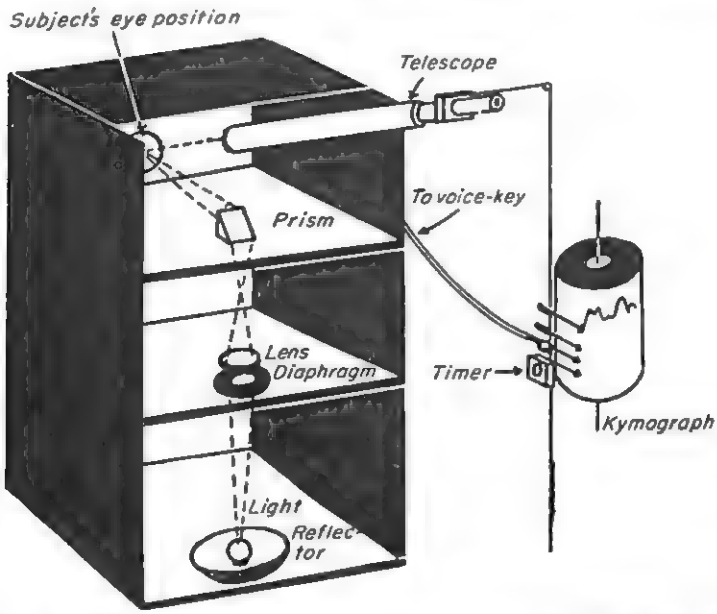


FIG. 99. A schema to show the arrangements used by Hudgins to study pupillary conditioning.

hair but also the entire telescope could be moved from side to side by means of a screw mounting. This adjustment enabled the experimenter to bring the hairs in proper position for measuring the pupil. The horizontal diameter of the pupil was measured directly from the micrometer scale in steps of 0.1 mm. After a little practice, pupil variations were quite easily and quickly followed.

The conditioning stimuli were bell sounds, contractions and relaxations of the hand, and spoken words. The bell was stationed in the cabinet slightly above the subject's line of regard. The hand contractions and relaxations were made in clasping a

standard dynamometer. The dynamometer was made to record contractions and relaxations electrically on a kymograph through contacts on the indicator and dial. A tension of 5 kg. was used for the male subjects and 4 kg. for the female. The circuits for the bell and light were designed to be closed either by the dynamometer or by a set of switches available to the experimenter. Switches in this position could also break the dynamometer circuits.

Figure 99 also schematically shows a kymograph for recording settings of the micrometer screw of the telescope, time from a Jacquet timer, the response of a voice-key, and the incidence of the conditioning stimuli.

The apparatus was placed in a darkroom. Fourteen subjects were employed, ranging from a staff member familiar with the problem to undergraduate students and one high school student, all entirely unaware of the nature and purpose of the study.

Procedure. The general design of the experiment involved the establishment of a conditioned response initially to a complex stimulus situation, which was later reduced to a simple form. The final objective was the substitution of a verbal process for the complex stimulus situation. The verbal process was to be effective, even when words were expressed subvocally.

To do this, the pupil of the eye was first conditioned to the sound of the bell, the light, of course, being the unconditioned stimulus. The subject's hand contraction and subsequent relaxation were next employed to close and open the light and bell circuits. Spoken words of the experimenter were used to elicit the subject's hand response. In the next phase of the procedure, the bell and hand reactions were eliminated. The remaining conditioned stimulus was the verbal expression used by the experimenter. Finally, the subject produced the verbal stimulus himself. First, the expression was spoken aloud, then in a whisper, and ultimately subvocally. Pupillary behavior was measured at every step of the procedure.

Results. Prior to training, the sound of the bell always elicited pupillary dilatation, as shown in Fig. 100. After conditioning, the pupil responded by constricting, as illustrated in Fig. 101. In general, the same result was found with hand contraction.

It evoked pupil dilatation. Only after a training period did hand contraction result in pupil constriction. The same thing happened in response to verbal stimuli. The pupil response was, at first, that of dilatation. After training it was constriction. This is shown in Fig. 102. The initial dilatation and subsequent constriction occurred both for verbal stimuli given by the experimenter and later for the verbal stimuli given by the subject himself (autoverbal stimuli).

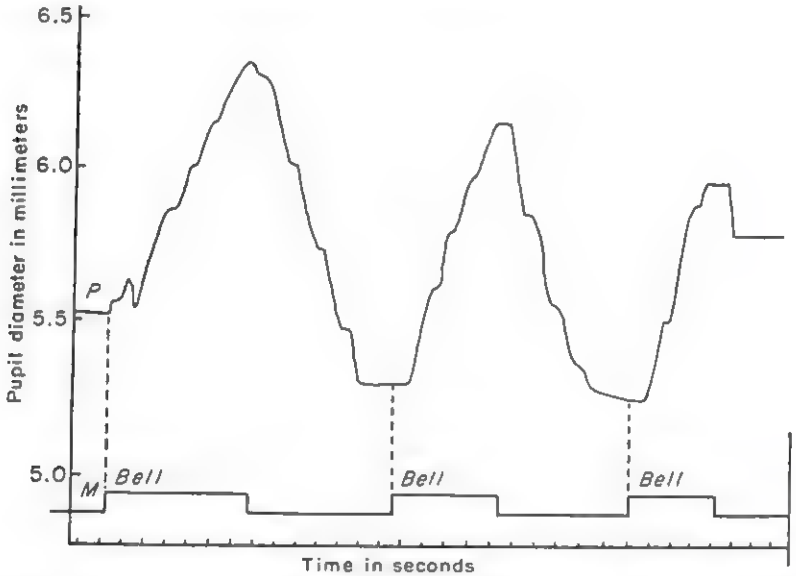


FIG. 100. Pupillary response to the sound of a bell prior to the conditioning process. (*Hudgins. J. gen. Psychol., Journal Press.*)

In summarizing, it may be said that words can become conditioned stimuli for the pupillary response. Conditioned pupillary responses can be evoked by words spoken by the experimenter or by the subject himself. The autoverbal stimuli are effective also when whispered or produced subvocally.

The conditioned responses to the verbal stimuli were relatively permanent. They were still present at least fifteen days after establishment. The latencies of the conditioned responses were from five to ten times as long as for the response to the light (unconditioned light reflex). The conditioning occurred in spite of the fact that the majority of the subjects were unaware that pupil responses were being conditioned or were involved at all.

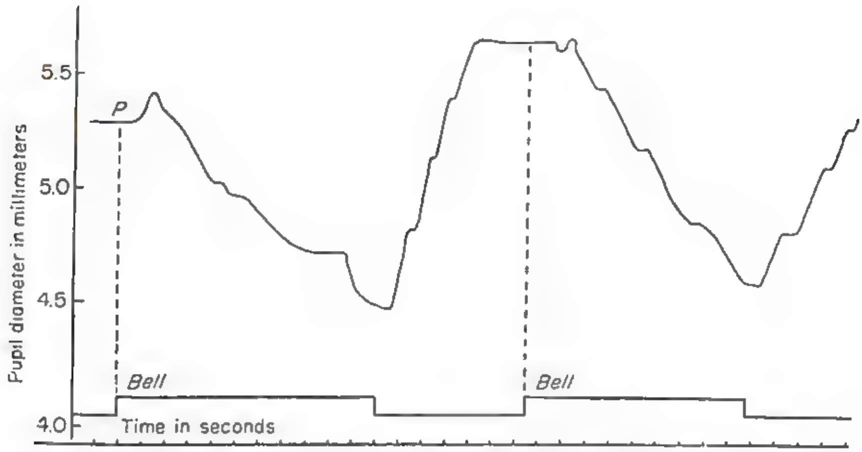


FIG. 101. Pupillary response to sound of bell following conditioning. Now the response is that of constriction, whereas to begin with it was dilatation. (*Hudgins. J. gen. Psychol., Journal Press.*)

Since some of the subjects did not know that their pupils were responding as they spoke the stimulus word "contract," it cannot be said that the conditioned pupillary response was a voluntary affair in the fullest sense of the word. On the other hand, Hudgins was led to express the hypothesis that so-called voluntary

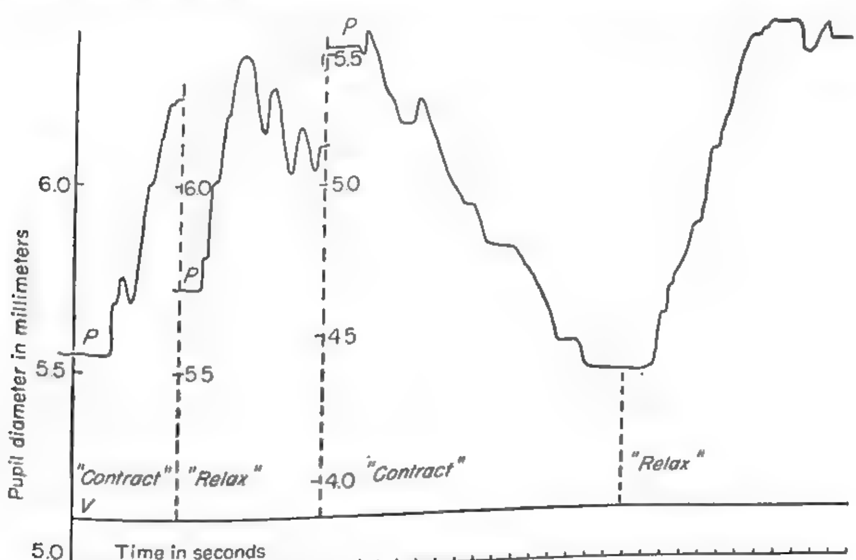


FIG. 102. Pupillary responses to spoken words "contract," and "relax," before (left) and after (right) conditioning. (*Hudgins. J. gen. Psychol., Journal Press.*)

behavior is of the very kind he demonstrated. That is, that voluntary behavior is the kind that is under the control of self-excited receptor processes.

To the naïve, that which is voluntary is that which happens immediately following a wish. It is merely the result of wanting to. The man-in-the-street would class walking, etc., as voluntary behavior, since it appears to be the immediate result of wanting to walk. Without elaborate experimental analysis such as is illustrated in the present experiment, he is unaware of the mechanisms involved in voluntary behavior and is bound to overlook the intervening processes between "wanting to" and actual execution. For that reason, he is not going to see any similarity between the evoking of pupillary responses and walking. The problem of how close the similarity may be is still to be discovered.

Another example of a shift from an involuntary reaction to the same reaction voluntarily induced is the following. If an individual looks at a nearby object, his two eyes converge or point toward it. For instance, if a pencil is held before the eyes and brought nearer and nearer to the face, the eyes converge more and more. This convergence may become so marked that the individual has the appearance of being cross-eyed. All of this happens even though the individual may be unaware of it. All that he is trying to do is see the pencil clearly.

Some individuals are able, at will, to turn their eyes toward each other (converge them). This they do without looking at anything in particular. In this respect it can be said that no specific external event or object is the stimulus. Even so, they are not necessarily able to converge their eyes by merely "wanting to." They may have to imagine that they are looking at something very close to their eyes or they may have to imagine how the eyes feel when they are greatly converged. This amounts to saying that there are certain intervening processes necessary. It may be that during early life when certain so-called voluntary performances were being learned, certain intervening processes were also necessary. This is a question that needs to be clarified. Even if the intervening processes were then necessary, they do not seem to be later on. It is from this absence of the performer's

realization (awareness) of any such intervening process that the conventional conviction of *voluntariness* arises with regard to many human performances.

QUESTIONS

1. Name the kinds of learning listed in this chapter.
2. What problem was Hudgins trying to solve?
3. Describe the apparatus Hudgins used.
4. What were the "conditioned stimuli" used?
5. What was the "unconditioned response" used?
6. Was the conditioned response the same as the unconditioned response? Indicate by telling what each was.
7. Indicate the sequence of steps used by Hudgins in arriving at his ultimate goal.
8. Were the conditioned responses to verbal stimuli relatively permanent, or quite transient?
9. Compare the latencies of the conditioned and unconditioned responses.
10. Summarize Hudgins' results, giving what you believe to be their significance.

CHAPTER 35

AN INVESTIGATION ON THE THEORY OF CONDITIONING

The various phenomena of conditioning elicit attention largely because they constitute the central experimental basis for a current type of psychological theory—associationistic behaviorism. Although the phenomena of conditioning prove interesting in themselves, their main service to psychology is the light they throw on fundamental processes of change in the organism. As concrete in its implications as conditioning seems to be, theory plays a considerable role in the conduct of experiments upon the subject. Men have to "make sense" out of the facts that come to their attention. They must invent causes or reasons why things happen as they do. In this, they are theorizing. They clothe events with theory. In turn, they clothe theory with so-called facts.

Zener's investigation of conditioning, which is to be described in this chapter, is one that contributes evidence that must be taken into account in any specific conditioning theory.

Few radical departures from Pavlovian theory have been attempted. Pavlov's ideas were based upon the idea of reflex response. Having begun with salivary reflexes, he thought in terms of reflexes even when he dealt with the most complex responses. They were essentially combinations of reflexes or irradiations of neural activity into collateral channels without any over-all guiding considerations. His was an associationistic and building-block theory. The rejection of his strict building-block theories did not actually lead most men beyond some form of associationism. There are exceptions to this rule, however.

Zener, among others, has felt that the simple associationistic idea regarding behavior does not describe what actually happens

in conditioning. The usual theories of the *substitute-stimulus* type assume that the change occurring in conditioning is the establishment of an association, connection, or tendency to excitation between the *conditioning-stimulus* system and a *response*. They assume that the response conditioned is substantially identical with the response that occurred during learning. This is considered obvious by most experimenters in this field. It is likewise demanded by the theories which they construct.

Notwithstanding, discrepancies occur between the unconditioned behavior and the conditioned. The ways in which the two responses are *not* identical have seldom been recognized in theorizing. Zener's evaluation of the matter led to the conclusion that the conditioned and unconditioned responses so often fail to be identical that the substitute-stimulus theory is clearly inadequate as a general conditioning theory. To take care of this defect, some have turned to kinds of so-called sign-urge theory. Zener was one who did this. Accordingly, we shall deal with the theory he advanced and his investigation designed to test it.

According to sign-urge theory (1) the essential modification on the animal's part consists in a reorganization whereby a kind of functional whole emerges from the mechanisms involved in perceiving the conditioned and unconditioned stimuli and the organism's urge or tension system that was initially excited by the unconditioned stimulus. (2) It assumes that the reorganized perceptual system activated by the conditioned stimulus alone sets up expectation of the unconditioned stimulus. The result of this is motivated behavior directed toward the signified goal. (3) Once the response is conditioned, the *intensity* and *persistence* of the behavior will vary primarily with the state of the urge system and with factors in the situation such as those pertaining to the perceived goal objects. (4) The direction of the behavior will be determined by the structure of the environment-organism field (as inferred, perceived, or expected by the organism). (5) The complex behavior (from any standpoint) is not fruitfully analyzable into component response elements or into a series of definable separate reactions. Behavior is directly

determined by the organization of the environment-organism field. (6) The most significant aspects of the field for determining the conditioned behavior include the sign-significance (goal) organization, means- and barrier-objects, behavioral supports, and the relations between all these and the urge or tension system of the organism.

It is hoped that this statement of what is involved in the sign-urge theory of behavior is not too complex and involved to be comprehended. As we progress with Zener's investigation, the issues in point should certainly become apparent.

Zener has several things to say about methods of experimentation in conditioning. He points out that the relative inadequacy of much of the usual experimental information to differentiate conclusively between alternate theories of conditioning is largely because the methods, by their nature, reflect the substitute-stimulus conception held by Pavlov and others of the original Russian school. The customary selections of simple reflexlike responses for investigation and of graphic methods of recording have demanded a higher degree of constraint of the animal subject than is found in other learning methods. Accordingly, the behavior that is permitted to occur outside of the selected response is extremely limited. It is often so slight that little attention is given to it, and any interpretation of it is highly ambiguous. Zener says that this is especially true of the factors involved in determining what *direction* the behavior takes. Behavioral directions can be accurately determined only by permitting trains of behavior to complete themselves.

Zener's methodological aims were to obtain as complete a description as possible of the outward behavior of the subjects during the manifestation of conditioned salivary responses set up by the conventional Pavlovian technique. He set out to elicit behavior both under conditions of training and under later systematic variations. Zener believed that no amount of behavior manifested under conditions identical with those of training could provide the investigator with the necessary knowledge of the nature of the changes in the animal that occurred during learning. Confining one's observations to the learning period alone is not likely to be productive of crucial evidence that

would differentiate between alternative conceptions of these changes.

Method and Procedure. No numerical records were made, but the animal was photographed so as to include all the overt behavior that was manifested. The animal was placed on a training table in a comparatively soundproof room located within a larger room from which the observations and recording and stimulus-regulating instruments were operated. A one-way vision mirror provided continuous observation of the animal's behavior from the outer room.

The conditioned stimulus in the investigation was a bell, producing a clang of about 65 db. in intensity, which was placed on a wall of the soundproof room. The bell was about 5 ft. in front of the animal and about 2½ ft. to the right of its head, slightly above head level. All the animals localized the source of the sound quite accurately. The food used for reinforcement consisted of four to six small pieces of dog biscuit put in a pan just below the animal's nose. The acid for reinforcement was 4 cc. of 0.25 per cent hydrochloric and was put into the animal's mouth through a metal tube, connected with a source of supply in the outer room. Salivary secretion was indicated by displacement of water in a tube system leading from a disk placed over the external opening of the parotid duct on the dog's cheek. This displacement was observed in a capillary tube in the outer chamber.

To condition the animals, the customary Pavlovian training procedure was used. The conditioned stimulus initially preceded the unconditioned by an interval of 1 sec. This interval was progressively lengthened to a 15-sec. delay. The conditioned stimulus (the clang of the bell) lasted for 25 sec. The food was presented at the intervals of delay already indicated. The behavior of the animal that we shall describe was manifested during the delay period. The experiments were conducted in the morning after 20 to 25 hr. of food deprivation.

Results. Various dogs behaved differently, and simple generalizations, therefore, would not be adequate to describe what they did. Zener, however, assures us that the actual range of behavioral responses manifested was far less than the total number

physically and physiologically possible. Certain dogs manifested certain tendencies to one kind of behavior, and other dogs to others, whereas some dogs behaved in a variety of ways.

Before conditioning, the experimental animals all behaved much more nearly alike in their responses to the bell and to the food than after conditioning had occurred. Typically, when the bell sounded, the animal turned quickly and looked at it. As the food mechanically dropped into the pan below the animal's nose, it lowered its head and ate one or more of the biscuits. From then on it continued to eat the food, either holding its head down in the pan or raising it between gulps.

After conditioning, the animal showed more variation in its behavior. It might look either at the bell or the food pan during the whole delay period. Some of the older and long-trained animals usually looked at the bell immediately and continued to gaze at it until the food dropped, or until very slightly before. Other dogs initially glanced at the bell and then looked at the opening of the chute where the food would appear, keeping this fixation throughout the delay period. This was characteristic of the apparently more hungry dogs. Often such dogs even omitted the initial glance at the bell. In some instances the dog kept its nose almost touching the bottom of the pan. In others, it alternated quite frequently between looking at the bell and the food pan. There were sometimes six or more glances back and forth from bell to food pan. Even this behavior varied. In some cases it appeared to be smooth and calculated; in others it was jerky and "hysterical" in appearance. The rate of this alternation in glance often increased as the foreperiod progressed. Although still other variations in the manner of behavior were pointed out by Zener, those already mentioned well indicate that activity, after conditioning, is somewhat varied. We need, however, to point out that not all of the behavioral variations in the animals were confined to head movement. Changes in bodily orientation and posture also occurred. If the animal was not at first in a good position to eat from the food pan, a general bodily readjustment to gain such a position occurred before the food fell. An example of general bodily orientation is the following: The animals in some cases made a slight but nonetheless definite

movement toward the bell as it sounded. This was followed by a later backing up so as to be in a better position to eat.

Zener observed responses in the animals also when they were removed from experimental restraint. For example, four animals were tested for a few days while unstrapped on the training table. In no instance during this time did any of the four dogs move away from the food pan when the stimulation was being given. Their performance was indistinguishable from that under the usual restraint, except for an enhancement of the initial slight motion toward the bell in the case of one animal. This applies to the instances in which food was used as the reinforcement (the unconditioned stimulus). In the experiments in which the acid was used for reinforcement, avoidance behavior became pronounced. In none of the trials did the animals increase in their approach movements toward the food pan. In some trials no gross change was observable, but in others (with three of the four dogs) the animal walked definitely away from the food pan before the full 15-sec. delay period had elapsed. Some of the animals even left the table, either by jumping off or by descending the stairs that led up to the table.

Discussion. It may be asked how these and other observations of Zener help to distinguish between various conditioning theories. As was indicated at the beginning of the chapter, the more frequent form of conditioning theory implies that the basic modification in the animal's behavior is the establishment of some association, connection, or tendency-to-respond between the condition stimulus and the response elicited by an unconditioned stimulus. It is taken for granted that this response (the conditioned one) is virtually identical with the response that occurred during the learning. It was because Zener, among others, believed from his observations that this idea did not fit the facts that he developed a concept that would recognize them and would postulate a basis for such facts. From his observations in the present investigation, Zener was able to say that, except for the simple salivary secretion itself, the conditioned and the unconditioned behaviors of the experimental animals were far from being identical. The "complex of movements" manifested in response to the unconditioned stimulus (food or acid) was not

the behavior that occurred to the conditioned stimulus. He believes that the variations in behavior that were observed are consistent only with some sort of a sign-urge conception—behavior pertinent not to the actual presence of the food in the pan, but rather to an expectation that it will soon appear. Both the so-called restless movements and the more directed and deliberate ones are in line with the sign-urge supposition.

Behavior When Conditions Are Changed. Another salient factor that Zener set out to deal with was the problem of what would happen were certain conditions in the general experiment manipulated. Some of these changes pertain to internal conditions in the animal, such as variations in the degree of hunger or of satiation. Other conditions involved placing the animal in new orientations with reference to the goal-object (food). We may briefly summarize the results of these two kinds of manipulations of conditions. Aside from the decrease in the salivation manifested with decreased hunger, there was a change in the overt behavior of the animal consisting in decreased alertness, constancy, or precision of fixation, and general muscular tenseness. There were also other more gross differences brought about by decreasing hunger. There was often a delay in looking into the food pan and a diminution or absence of the alternate glancing at the food pan and the bell. Often the bell was not looked at at all. Few restless movements appeared. As hunger decreased the animal might not look into the pan until food appeared, and sometimes not even then.

If several types of food were used for reinforcement (unconditioned stimuli), salivation secretion tended, in general, to decrease as a result of the food's random or alternate presentation. Whereas, during satiation, there was a change in the conditioned response (overt behavior and salivary secretion) when reinforced by food, this change was absent when reinforcement with acid was used.

Such an outcome, while completely within the requirements of the sign-urge hypothesis, is fully incompatible with the traditional substitute-stimulus hypothesis. Zener believes that the facts taken as a whole indicate that the entire idea implied by the term "conditioned response" is inappropriate. The facts

very strongly suggest that "a" response, even a very complicated one, is not conditioned as is usually believed. The change called "conditioning" that occurs in the animal is not a structural change, such as the production of a connection between the conditioned stimulus and a definite conditioned response, but rather a reorganization of the animal's perceptual system in relation to its system of urges, the combined action of which results in not a highly specific response but in kinds of behavior directed toward given goals. Entering into the total picture of the overall situation, aside from what we have just mentioned, is the organization pertaining to means to the end and the barriers that exist in the given situation. One of the chief reasons why conditioned response so often resembles the unconditioned as much as it does is that the experimental limitations are usually kept quite nearly identical in training and test situations.

QUESTIONS

1. What is the main service to psychology of conditioning experiments?
2. What descriptive label does Zener apply to the usual type of conditioning theories?
3. What do these theories imply?
4. What is Zener's chief criticism of the usual conditioning theories?
5. What label has Zener given to the type of theory of which his own is an example, and what are its essentials?
6. What must be done in addition to studying the learning period itself?
7. Describe Zener's experimental procedure.
8. Describe the behavior of his animals in the training experiments.
9. Describe the animals' behavior when conditions were manipulated.
10. What did Zener conclude from his experiments?

CHAPTER 36

INTERPOLATED LEARNING AND RETROACTIVE INHIBITION

One of the salient features of the learning-forgetting process is *retroactive inhibition.* Retroactive inhibition is a principle used to account for forgetting as a consequence of learning new material interpolated between the memorization and later recall of the initial material.

For those interested in problems of learning, the functional relationship between the degree of learning and the extent of retroactive inhibition is an important matter. McGeoch pointed out that it is not only necessary to know this relationship (where inhibition is demonstrated) but also to know whether a degree of learning may be reached which puts the material beyond the reach of inhibition. McGeoch had previously shown that the relative amount of retroactive inhibition varies inversely with the number of presentations given of the lists of learning material. In his study, the interpolated material was kept constant at 11 repetitions. The repetitions of the initial learning material were varied in steps of 5, from 6 to 26. Even when the repetitions reached the maximum (26) and where some "overlearning" was involved, the amount of inhibition was 46.2 per cent in terms of the recall score and 5.3 per cent in terms of the saving score. This indicated that a considerable amount of overlearning did not rid the material (nonsense syllables) from the detrimental effects of interpolated learning.

These results led McGeoch to the question of what would happen were the interpolated material instead of the original lists varied in degree of learning.

For the reader who is not familiar with the concept of overlearning, it is pointed out here that in experimentation on

learning, a criterion is always set up to indicate when material is said to be learned. For example, the criterion chosen may be that of being able to respond correctly in each of two successive stimulus presentations. Any additional learning trials beyond this are considered productive of overlearning. Consequently overlearning is taken to be all learning beyond a certain minimal degree of mastery and is a purely arbitrary matter.

McGeoch's Study. For the present chapter, we shall devote attention to a study by McGeoch such as already suggested. The problem in this study was to measure the influence of varying amounts of interpolated learning upon amount of retroactive inhibition. The conditions were to be comparable to those in the earlier study already mentioned. The original learning lists were to be presented a constant number of times, the assumption being that the degree of learning was more or less constant when the repetitions were constant.

Method. Two separate experiments were conducted. For the first, the same lists of nine nonsense syllables as used in the earlier study were employed. They were exposed at a 2-sec. rate on a drum and were learned by the procedure of *anticipation*. In this, the subjects were required to anticipate each syllable by spelling it orally rather than by merely pronouncing it. Since the initial syllable of the nine was a cue for the succeeding one, only eight syllables actually constituted the learning material.

The original lists were presented eleven times, and the interpolated lists were presented 6, 11, 16, 21, or 26 times. Six experimental sessions were required, with one rest condition and one work condition for each of the five frequencies of interpolation used.

For the rest condition, the subject was given 11 presentations of the original material and then a 10-min. period of so-called nonlearning activity. During this period he was given a book of jokes to read and from which to select the three best out of every ten. This task was given the subject in the attempt to prevent his rehearsal of the learning material just previously presented.

For the work condition, after the subject was given the 11 trials of the original material, he was immediately given either 6, 11, 16, 21, or 26 presentations of an interpolation list. The unfilled

remainder of the 10-min. period was occupied by the same activity as in the rest condition. At the end of the 10-min. period, the original list was presented again and repeated until the subject could anticipate each syllable correctly in three consecutive runs. After this, the interpolation list was presented again to ensure the idea that it, too, was important in the experiment.

For the foregoing procedures, 12 college students (7 females; 5 males) were used. Each of them went through the whole sequence of rest and five work periods twice (two cycles). This they did in an order which counterbalanced positions of syllable lists and practice effects.

Results. The first determination in the examination of the results was whether the constant number of repetitions used in the learning of the original lists was accompanied by a constancy in the number of correct syllable anticipations. The number of correct *anticipations* was taken as a control measure of whether the conditions in the learning process actually resulted in a uniformity that might rightly be called a constant degree of learning. McGeoch found no significant differences in the number of anticipations when rest was interpolated between learning and retest, and when interpolated lists were repeated 6, 11, 16, 21, or 26 times during the interim. From such a result, he assumed that differences in amounts of retroactive inhibition developed under the different amounts of interpolation are not altered to any significant extent by possible inequalities in the degree of learning of the original lists.

The correct recalls were also examined. These were represented by the numbers of syllables correctly anticipated on the *first* relearning trial. All of the interpolation frequencies result in considerable inhibition, but the amount produced by the six repetitions was from 19 to 26 per cent less than produced by the other frequencies. Frequencies from 11 to 26 seemed to produce about equal amounts of inhibition. The third set of results, those which pertained to relearning to a criterion of three perfect trials, was also examined. Approximately half the original trials were saved when relearning was begun after interpolated rest. Far less than half were saved for the relearning series begun after interpolated lists of 6 to 26 trials. Here again, six

repetitions of the interpolated lists were less effective (produced less inhibition) than any of the greater number of repetitions. McGeoch concluded that when the interpolated learning is little, retroaction seems to be a function of the degree of learning, but as complete mastery is reached and passed, additional increments have no noticeable effect.

McGeoch also determined whether the number of repetitions of the *interpolated* lists of syllables was accompanied by a proportionate increase in the number of correct anticipations of the syllables in them. It was found that the number of correct anticipations increased in an almost linear fashion as frequency increased from 6 to 26 repetitions.

Furthermore, a comparison was made between certain results in the present investigation and in a previous one. Whereas in the present study, the variant was the number of repetitions of the interpolated material, in the previous one the variant was the frequency of repetition of the original lists being learned.

From the comparison it became clear that the influence of variation in frequency was a function of the particular list in the retroaction-arrangement to which it was applied. When a uniform frequency of 11 repetitions was used in the interpolated lists, and the frequency of presentation of the original lists was varied, retroaction decreased progressively as frequency was increased. In the present study, retroactive inhibition increased only as repetitions of the interpolated material was varied from 6 to 11, remaining constant from there on. Therefore, increasing the extent of learning of the original lists freed them more and more from the opposing effects of constant amount of interpolation. Beyond the few repetitions, increasing the number of repetitions of the interpolated material had no enhancing or diminishing effect upon the original lists learned. It was especially significant that retroaction, as measured by relearning, remained marked in the latter case and decreased decidedly in the former.

The amount of practice under the conditions in which retroaction was measured was another factor that apparently influenced the amount of retroaction occurring. This factor was examined by McGeoch in the present experiment by comparing the amount of retroactive inhibition occurring in each of two

cycles of learning. It was found that retroaction, in terms of recall score, decreased decidedly from the first to the second cycle for the six-repetition interpolation only. Measured in terms of relearning score, there was a very large decrease in retroactive inhibition in the second cycle. In this cycle, there was much less inhibition in relearning after a 26-repetition interpolation than for any of the other frequencies beyond six. Although the difference was not statistically reliable, it suggested to McGeoch that were practice continued far enough, there might have been an inverse relation between frequency and inhibition, at least at the higher frequencies.

The influence of practice is, of course, a function of the way retention is measured. The influence is almost zero when measured by recall and very decided in relearning tests.

For the second experiment in the investigation, McGeoch used 10 nonsense syllables that were given to the subjects for 120 sec. to study. The time given to study interpolation lists varied from 60 to 180 sec. in steps of 30 sec. The subjects were given 45 sec. in which to write their recalled list of syllables. After this they were given 5 min. of "rest," as in the first experiment. At the end of the rest period, 45 sec. were allowed for a second recall.

For the interpolated "work" conditions the subjects were given a second list of syllables to learn after the initial 45-sec. recall period for the first list. For the interpolated list, varying amounts of time, as was already pointed out, were used. At the termination of the interpolated work period, 45 sec. for a second recall of the lists just studied were given. The part of the 5-min. period that still remained was used in reading the jokes as during the rest condition. At the end of this period, 45 sec. were given for a second recall of the original lists. A second recall of the interpolated lists were required so as to remove suspicion that these lists were unimportant. Forty-two college students were used for this experiment, in groups of three, with one rest and five work conditions.

The data were scored in two different ways: (1) by giving a score of 1 to each syllable recalled correctly and in its right position in the list; and (2) by assigning partial credit in terms of both syllable correctness and right position in the list.

To obtain a measure of retroactive inhibition, the number of syllables recalled after the interval was divided by the number recalled originally. The percentage of retroactive inhibition was the percentage recalled after rest, minus the percentage recalled after a given work condition divided by the percentage after rest or $(R-W)/R$.

The scores obtained for the first recall under the one rest and five work conditions (*i.e.*, prior to rest or work) fluctuate somewhat, but not enough to invalidate the idea of a constant amount of learning having taken place. The second recall (after rest) was definitely greater than the recalls under any of the work conditions. The difference was statistically reliable. The differences between recalls after the interpolated work periods were only slight. Due to the amounts of decrease in recall as the subjects varied from 60 to 150 sec. of interpolated work, the suggestion was that retroactive inhibition increases slowly with increase in time spent in interpolated learning. There was a slight reversal in the data regarding 150 and 180 sec., so that it is possible that retroactive inhibition tapers off or actually becomes less with more extended learning of the interpolated material.

At least it may be said that retroactive inhibition reduces the retention of material considerably, even with as little as 60 sec. of interpolated study and with as much as 240 sec. of it.

McGeoch, in a discussion of the possible factors inherent in the concept of "degree of learning," pointed out that it is not a simple affair, but rather that there are several kinds of "degrees of learning." These must be recognized and reckoned with in investigations on retroactive inhibition and in retention in general.

McGeoch asserted that the usual methods for denoting the degree of learning by no means exhaust the possibilities. Even in apparently simple experimental procedures there are uncontrolled variables whose possible influence is entirely undetermined.

It is inevitable that when frequency of presentation of items is varied, the time spent in learning them varies as a consequence. A given number of correct anticipations (or performance to any other criterion) might be achieved by different subjects after

greatly different numbers of trials, and therefore after different durations of time. It is conceivable that the variant of time might, in itself, despite a constant number of correct anticipations, lead to varied results.

It is conceivable that when learning is imperfect the relation between the per cent learned and per cent to be learned may be a factor.

Measurement in terms of numbers of correct anticipations is, at best, only approximate in what it signifies, since the distribution of the correct anticipations over the list used is not taken into account. It is an additional possibility that the portion of the learning period in which an item is first mastered may be a factor. This factor would be a function of others, such as the homogeneity of the list, serial position of the item, etc.

From such considerations as these, McGeoch felt that it should be apparent that the problem as to what constitutes degree of learning and its influence upon retroactive inhibition have been only barely touched upon experimentally. Although the results that he obtained were valid under the conditions used in the experiments, the conditions represent only a very few among many.

QUESTIONS

1. What is retroactive inhibition?
2. What is one variable that may be related to the degree or extent of retroactive inhibition?
3. What was McGeoch's problem in the investigation presented in this chapter?
4. Describe McGeoch's first experiment.
5. What was the first matter to be considered in examining the results of this experiment?
6. What were McGeoch's findings in this respect?
7. What else was determined in the first experiment?
8. Describe McGeoch's second experiment.
9. How were the data in this experiment scored?
10. What had McGeoch to say about degrees of learning?

CHAPTER 37

SOME PROPERTIES OF VISUAL RECALL

In the study of many phases of learning, recall is used as a criterion. This is to say that to study learning, memory is very often used. Despite the frequent involvement of memory in this way, memory itself is very seldom studied. This omission of the study of memory has been true from the very first. Even Ebbinghaus' classical study of memory was primarily concerned with learning and not mainly with memory.

In this chapter, it is our intention to deal with memory. For this purpose we have chosen Perkins' study on the symmetry of visual recall. Perkins seemed to be very vividly aware of the distinction between studying learning and memory and laid down several rules that should be followed in memory studies. For example, he said that in order to carry out an investigation on memory, the material should be presented to the subject only once. This stipulation was, of course, not carried out by most investigators. The material should be simple enough for the single presentation to be sufficient. This is not so obviously necessary as was the first rule, for it is not obvious why, in all memory studies, the perception and the first reproduction of the material should be complete or correct. Such a starting point as a comparison for later reproductions is only one of the many possible starting points. Perkins' third rule is that there should be only a few stimulus figures, so that the subject will be able to reproduce them all after the single presentation. This rests largely upon the same assumptions as the second rule. It is at least a practical rule, if not a theoretical necessity. The fourth rule states that the material should be reproduced entirely from memory at various time intervals, so that the trend, direction, and extent of the changes as time passes will be made observable.

The picture of the memory process will, accordingly, be disclosed under such conditions as these.

We agree to the distinction between memory and learning implied in Perkins' criteria and work. Yet it would seem that as far as the degree of difficulty of the memory material is concerned, it should be varied from the most simple to the very complex. Surely discoveries regarding memory would be possible under this variation that might be by-passed by omitting it. This variation might help to reconcile the results and interpretations of such studies as the following.

Phillippe, who in 1897 was the first worker to use the method of presenting figures and requiring the subject later to make drawings of them, classified his results as follows: (1) Detail becomes lost or grows vague and confused — "useless" details disappear; (2) new details replace the old; and (3) the material takes on a generalized form and approaches the typical form of certain objects.

Kuhlmann, in 1906, came to the conclusion that many errors of reproduction came about by dependence upon verbalizations used initially to describe the figures. Verbalizations were influential in determining results also in the experiments of Carmichael, Hogan, and Walters. The errors, according to Kuhlmann, seemed to be largely of two kinds: (1) Object assimilation. The subject remembered the object that he perceived the original to resemble. The reproduction (drawing) was distorted in that direction. (2) Several kinds of "regularization," or distortion in the direction of certain familiar geometrical forms. Parts of figures off the vertical or horizontal were made vertical or horizontal. Parts not quite parallel or perpendicular to other parts tended to be made parallel or perpendicular. Parts not quite equally long were made so. Asymmetrical parts tended to be arranged symmetrically.

Meyer, in 1913, using six named figures presented them several times before requiring drawings 24 hr. later. The errors that resulted were classified into three categories: (1) reproducing figure as a whole with omission of a number of details; (2) object assimilation; and (3) the confusion of figures within the series.

Bartlett, in 1932, likewise conducted an experiment of the same general kind that we are listing. Among the changes (errors) he found were (1) the enhancement of unusual details; (2) increase in the complexity in some cases; and (3) reversals from side to side in asymmetrical figures.

Wulf, in 1922, under the impetus of configuration theory, postulated that a reproduced figure should be more "pregnant" (more like "good" forms) than the originally seen figure and that changes in this direction should be progressive with elapse of time. He categorized his results into three headings: *normalization*, *emphasis*, and *inherent structural change*. Normalization in a figure is deviation from the original in the direction of a familiar object. In other words, Wulf found "object assimilation." Emphasis is a change toward accentuation of some aspect of the original. Inherent structural change is an alteration of a figure such as some configurationists would say represents the yielding to inherent stress in its form. These are highly interpretative ways of stating matters and are difficult to follow.

Still other studies in the reproduction of visually perceived forms have been made in the past several decades. What is most obvious is that the conditions set up for presentation and recall are quite different from study to study. Some of the results are identical in almost all of them; but others, taken as they stand (and without due allowance for differences in experimental conditions), seem to be somewhat discordant. This leaves room for some new investigator to arrange an experimental design so as to tie all the results together into a consistent whole.

In this connection, one must recognize the possibility of variations of interpretation stemming from differences in schooling of the workers themselves, regardless of factual outcome. In this connection, we are reminded of what Kuhlmannn asserted. He said that "recall" is quite largely not recall at all, and that what the subject does cannot accurately be called reproduction of the original figure. Kuhlmann asserts that what the subject does is to *construct*, not reconstruct. Memory is thus only one of the several factors involved in "drawing something from memory."

Perkins' Experiment. This experiment was conducted for the purpose of determining the nature of the changes that occur in visually perceived objects as time passes. This was about the same problem as that involved in the studies of Wulf and of Gibson. The material differed, however. Two sets of material, *A* and *B*, were used. The method involved in each was the same. The two sets of five simple asymmetrical figures are shown in Fig. 103. Each figure was drawn on a separate card, 11 by 14 in. Cards *A* are shown in the top row and cards *B* in the bottom. Different subjects were used for the two sets of cards. A total of 150 subjects, 98 for cards *A*, and 52 for cards *B*, were employed.

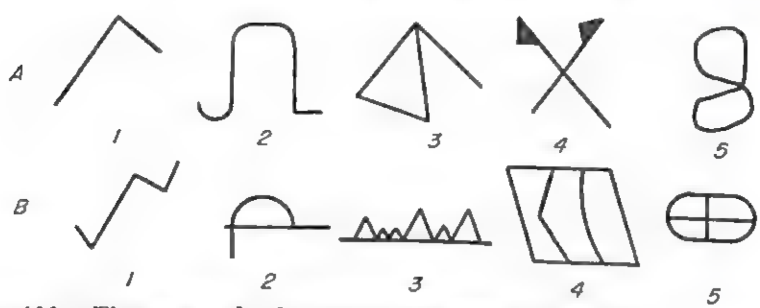


FIG. 103. Figures used for study of memory. (*Perkins. Amer. J. Psychol.*, 1932, **44**, 474, *Fig.* 1.)

All the subjects were students of elementary psychology, and none of them had any knowledge of the purpose of the experiment. The cards were presented to about 20 subjects at a time. The only instructions were to look at the group of drawings. Each card was then exposed for 5 sec. This, of course, was the only time the cards were presented. Twenty seconds after the exposure series was complete, the subjects were required to draw the figures as accurately as they could. After the elapse of 2 days, the subjects were required to reproduce the figures again. This process was repeated again on the third, ninth, sixteenth, thirtieth, and forty-ninth days following the original presentation. By this method, Perkins felt that a series of reproductions showing the undisturbed trend of change with successive additions of time was obtained.

Results. As a total, there were 3,559 "reproductions" of the 10 stimulus figures. In all but a few of these there were de-

viations from the original figures greater than would be expected from mere inability to draw freehanded. Perkins conceived his problem as being one of discovering general trends pertaining to the drawings and not that of classifying in detail the literal changes made. To begin with, the variations looked as diverse from one another as the subjects themselves.

In a few drawings, early in the experiment, no change from the original occurred, but as time passed changes began to occur. As

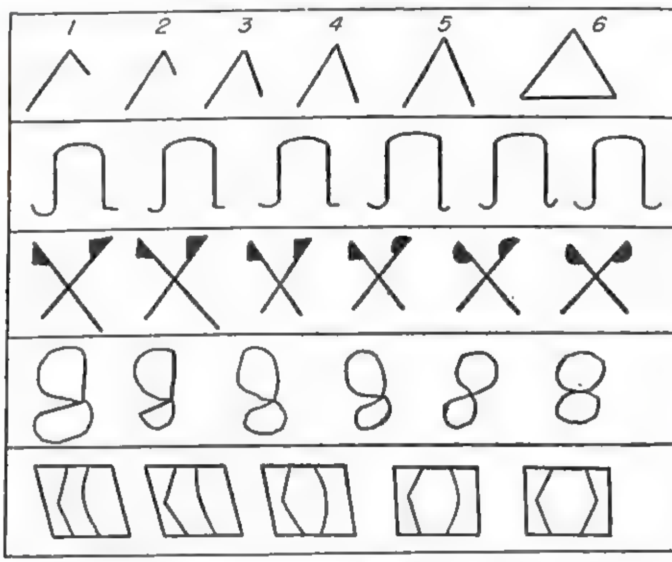


FIG. 104. Changes in reproductions of Perkins' original figures. (*From Perkins. Amer. J. Psychol.*, 1932, **44**, 482, *Fig.* 6.)

far as Perkins was concerned, all of the changes seemed to be toward symmetry (see Fig. 104). Changes in this direction for each of the 10 figures were tabulated. Coming within the category of increasing symmetry were nine general groups of changes, applying in most cases to all figures. These were given as follows:

1. *Equalization.* The making equal of those parts within a figure that were not originally equal. This equalization involved two features of the figures in the present experiment—equalization of lines and equalization of angles. Number 1 of Fig. 105 is an example in which equalization of both lines and angles occurred.

In some figures both could not have occurred, for the two would have been geometrically impossible.

2. *Orientation.* This occurred in one of three ways. It sometimes occurred as a verticalization of a part of a figure that lay originally in some other direction. (Fig. 105, No. 2.) At other times the same shift held true for the near-horizontal. At still other times there were reversals or inversions, as in No. 4.

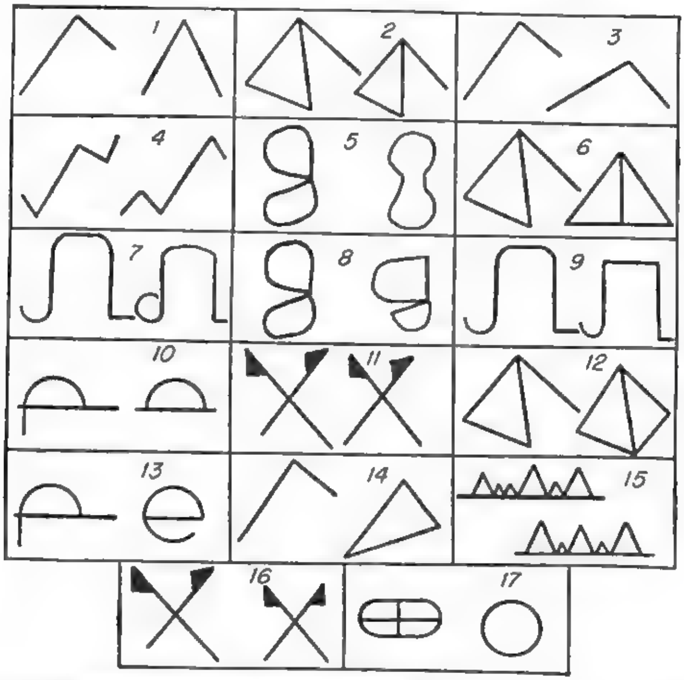


FIG. 105. Types of shifts toward greater symmetry in reproduction of visually perceived objects. (*Perkins. Amer. J. Psychol.*, 1932, **44**, 476, *Fig.* 2.)

3. *Standardization.* This is the shift toward the form of some well-known or standard object, such as a dumbbell or dollar sign, as in No. 5. This standardization may pertain to shifts toward standard geometrical figures, such as is illustrated in No. 6. Trends toward standardization also took the form of making parts into circles, ovals, or right angles, as in Nos. 7, 8, and 9.

4. *Simplification.* This involved the omission of parts or lines. These parts were those interpreted by Perkins to be destructive of the symmetry of the figures, as in No. 10. Simplification

was the name given to the shortening of parts to arrive at balance, as in No. 11.

5. *Complication.* This took place in two ways, either by addition of a part, or by increasing a part. Additions such as in Nos. 12 and 13 took place.

6. *Completion.* Numbers 12 and 14 would seem to be cases of this.

7. *Proportioning the relations.* This involved changing the position of some part of a figure, the part having no initial relation to its associates except position, so that the final proportions are better. This is shown in No. 15.

8. *Shifting toward bilateral symmetry*, as in No. 16.

9. *Shifting toward radial symmetry*, or as Perkins calls it, whole-symmetry. This is shown in No 17.

The changes that took place in the drawings were reckoned in terms of number. On this basis, the greater number of changes from the original took place in the first reproduction. Later changes of this sort were progressively fewer. Such changes were less often "new" ones. They were generally accentuations of changes already begun in the earlier reproductions.

Whereas the first reproduction represented the stage involving, in general, the greatest number of changes, it was the time when no changes occurred in some figures. The initial no-change figures generally underwent change later on, although only 1 per cent or less of the drawings manifested no change in the final reproduction.

Perkins found that at the end of the period (50 days), 40 per cent of the drawings made at that time showed perfect symmetry, whereas the percentage was only about 16 per cent at the end of the first 10 days. As in the case of general tendency toward change in the early part of the experimental period, the change toward perfect symmetry was proportionally more marked then than later. This was also true of change toward "semiperfect" symmetry.

Since not all of the observers made their reproductions according to the schedule listed in the earlier part of the chapter, varying amounts of time elapsed between their first and second reproductions, or between certain subsequent reproductions. In taking

this into account, it was found that in the recall of visually perceived forms the tendency toward symmetry is more pronounced when the time interval between them is long. This was true even though the total elapsed time for the reproductions was the same. The results might have been put in another way by saying that the tendency toward symmetry in a given period of elapsed time is greater the fewer the reproductions. Perkins does not state, in detail, what the number of cases and the intervals were, so that it is possible that it would have been better had the original experimental design systematically provided for varying amounts of time between reproductions and varying numbers of reproductions in a given over-all elapsed time.

Perkins concluded, from the evidence that he did have, that the drawing of the figures must, in itself, serve to oppose the tendency toward symmetry when spaced relatively closely in time. Despite this, it seemed that toward the last, the tendency toward change and toward symmetry emerged. From the progressive nature of the changes taking place from reproduction to reproduction, it was concluded that in some cases the changes might have been more pronounced had still more time been permitted. Thus, alterations representing only slight degrees of shift toward symmetry during the 50 days might have been more pronounced later on.

Perkins' results have considerable in common with those of other workers, particularly those of Gibson and Wulf; consequently he discussed the interpretations they gave to their results. Finally, he suggested a theory of his own. He said that an adequate theory, whatever it may be, must fulfill certain requirements in order to cope with the problems that exist. (1) The theory must be based upon the idea of mental continuity. Actually, memory and mental continuity are two aspects of the same problem. (2) The theory must be dynamic so as to explain the progressiveness of the alterations that occur in recall. (3) The theory must account for the direction of the alterations found in recall. (4) It must account for the greater changes that occur when the intervals between recall are long.

Perkins attempted to give a theory that possessed these qualifications. The theory was constructed in configurational and

organismic terms. In terms of the organismic school of thought, it was quite characteristic. This school has not been worked ardently enough as yet by its sponsors to gain an adequate hearing and to be properly appreciated.

QUESTIONS

1. Give Perkins' distinctions between studies of memory and of earning.
2. State some findings of earlier workers in the area of visual recall.
3. What was the purpose of Perkins' investigation?
4. Describe what Perkins did.
5. What kinds of changes were grouped together as examples of increasing symmetry?
6. What modifications or variations in Perkins' procedure could you suggest? Indicate what question you would intend to answer by them.
7. What effect was believed to result from the *drawing* of the figures?
8. Did the results of this investigation have much in common with those of others? If so, whose experiments, in particular?
9. What were the characteristics of a satisfactory theory regarding memory, as stated by Perkins?
10. In what terms was Perkins' theory stated?

PART VIII

PREPAREDNESS AND ADEQUACY

CHAPTER 38

THE FACTORS OF READINESS AND UNCERTAINTY IN RESPONSE

One of the aspects of human behavior that we have not yet touched upon pertains to the organism's own structure and orientation. To be more specific, we have not dealt with the properties of the organism's *directionality*, *continuity*, and *adequacy*. Actually, there are many terms that are used to connote what is meant in this area. Among them, for example, are attitude, posture, readiness, set, determining tendency, purpose, aspiration, and drive. These terms, although differing somewhat in their connotations and emphasis, greatly overlap. Naturally, it is impossible for us to do much toward covering such a field by describing a few investigations. Three chapters, however, will be devoted to it. The present chapter will deal with the problem of readiness largely as it is manifested in reaction-time experiments.

Reaction-time experiments are of many kinds and have long been a stock procedure in psychology. To measure reaction time, some of the earliest precision instruments in psychology were constructed. In essence, a reaction-time experiment is a formal arrangement in which the subject is instructed to respond in some way to the sudden appearance of a stimulus.. The stimulus may be a simple event that will have the same characteristics on each appearance, or it may be complex, differing in some particular from trial to trial, and requiring response in some cases and not in others. It may require one kind of response in one instance and different kinds in others. Generally, the response is a motor affair, much as the pressing of a key. It may be, however, a verbal response whose time of appearance is recorded by

a "voice-key." One of the earliest concepts of the reaction under these general conditions was that it was a "prepared reflex."

Forewarned stimulation involves *readiness* of some sort or other. Readiness is particularly involved in the requirement to react as rapidly as possible. In fact, since in most experiments of this kind it is essential to elicit the earliest possible response, a "forewarning signal," or "ready signal" is given prior to each appearance of the stimulus event. Our main interest lies in what the individual does to get ready to react and what must be done methodologically to ensure the necessary controls so that the results will not be falsely interpreted.

It was discovered that in reaction-time experiments the subject can become overready. This is evidenced by "jumping the gun." In such cases, the individual may react when the formal stimulus is not given, or he may react almost simultaneously with it. It thus became clear that all response times obtained in the usual routine way might not actually represent reaction time. On this account, something needed to be done to reduce, eliminate, or otherwise take into account the tendency to jump the gun. Checks called *Vexierversuche* (in German), or "catch-tests," were introduced. These were put into the series of trials in the attempt, on the experimenter's part, to train out the excessive readiness. The catch-tests consisted in giving the ready signal and then omitting the stimulus in certain randomly distributed trials. The tendency of the overready subject was to react in the absence of the stimulus. It was only after the subject ceased, or practically ceased, to make false reactions that reaction-time measurements were begun. Naturally, the false signals were not numerous, else the reaction experiment literally and fully become a choice-reaction experiment, and the length of the reaction times would increase accordingly.

The two major types of readiness in a reaction-time experiment are functions of the nature of the anticipation involved. One is the readiness to react, to do something, such as press a key. Reaction in this case is spoken of as a "muscular reaction." The other form of readiness is the readiness to detect the stimulus, the signal to react. This form is spoken of as "sensorial reaction." The muscular reaction occurs the sooner.

Reaction-time experiments were conducted in the early days to attempt to understand sensation, thought, etc. Visual, auditory, tactile, pain, and other types of stimuli were used. Auditory stimuli elicited the quickest response (shortest reaction time). Reaction to visual stimuli came next and was of about the same value as tactile reaction time.

Readiness was studied by varying the foreperiod, the period between the giving of the ready signal and the actual stimulus. Foreperiods varying in range from ½ sec. to at least 24 sec. were investigated. Foreperiods were kept uniform in length in some cases, and in others their length was varied in random fashion. One investigator used auditory stimuli given at intervals of ½, 1, 2, or 4 sec., in a continuous stream. Naturally, in making each response, the subject knew that another signal would appear shortly. Under such conditions the results were as follows: The mean reaction time for the ½-sec. foreperiods was 335 msec.; for the 1-sec. period, 241 msec.; for the 2-sec. period, 245 msec.; and for the 4-sec. period, 276 msec. The shorter the reaction time, the less variable was the response. The ranges of variability of the reaction times just mentioned were 64, 43, 51, and 56 msec. respectively.

Readiness varies in two respects: in its promptness, and in its duration. Some subjects get ready early and then lose this readiness, as attested to by their increased reaction times when the foreperiod is lengthened. Some subjects maintain their readiness for as long as 10 sec. In Woodrow's study of reaction times, the minimal reaction time tended to occur when the foreperiod was about 2 sec., and tended to rise as the foreperiod was lengthened. At the end of 24 sec., it was still rising for some subjects. There is some possibility that readiness may fluctuate during a 24-sec. period. This fluctuation would not necessarily be detected by the usual methods of experimenting. Evidence for a periodic fluctuation would, however, show up in the size of the variation in reaction time recorded. If these variations were found to be great, periodic fluctuation might be surmised.

There are two separate investigations which we wish to describe in the remaining part of this chapter. One is a study by Lemmon and Geisinger in which reaction times under light and under dark

adaptation were measured. The other is a study by Kellogg in which reaction times were measured as an adjunct to the ordinary procedures in a standard experiment in psychophysics.

Reaction in Light and Darkness. The study of Lemmon and Geisinger follows others in which it was demonstrated that the reaction time to peripheral visual stimuli was greater than to foveal stimuli. Most of these measurements were taken under conditions of light adaptation.

Lemmon and Geisinger used 14 college students as subjects. The reaction times were measured by a Dunlap chronoscope located in a room separate from the subjects. Forty-eight reactions were made by each subject under light adaptation and an equal number under dark adaptation. This gave a total of 624 reactions for the group under the conditions of light and 672 under dark adaptation. One of the 14 subjects did not serve for the light-adaptation experiments.

Procedure. The subject was seated with his head in a rest. The stimulus was a disk of light 1 cm. in diameter as seen from a distance of 58 cm. The visual angle subtended was thus roughly 1 deg., a value assuming stimulation within the fovea. The experimental room was lighted by two large overhead lights. White sheets were suspended about 2 m. in front of the subject, occupying virtually the total visual field. Two 200-watt lamps with white reflectors were located behind the subject. Thus a brightly illuminated field was provided. No photometric readings were taken, but it was certain that the level fell within an ordinary range for light adaptation. The stimulus disk was centered in a small white box bearing a Mazda flashlight bulb.

For peripheral stimulation, the subject fixated on a white pinhead on a white card 58 cm. in front of him. The stimulus disk was shifted in an arc to the left so as to make a 45-deg. angle with the line of regard and remaining 58 cm. from the eyes.

Both eyes were used in the observations. The ready signal was a buzzer, and an approximately 2-sec. foreperiod was given. The lamp, of course, did not light instantly, but since this was roughly a constant from trial to trial, the investigators did nothing about correcting the raw data. We might safely subtract 50 msec. from the readings which they offer. The light, when

once lit, remained lighted until the response was made, hence its duration was beyond the critical value.

For measurements under dark adaptation, the room was absolutely dark. The subject was adapted for 30 min. The box bearing the stimulus disk was painted black instead of white. The level of stimulus used was one in which the disk was just barely visible after the full adaptation period. The fixation point was a lighted pinhead illuminated by a faint projection from an invisible offside source. The procedure was the same as for the light-adapted reactions.

Results. All but one subject reacted more slowly to peripheral than to foveal stimulation, the average additional time being about 35 msec. The ratio of this value to the standard deviation was 6.78, showing a high reliability in the readings.

In dark adaptation, one-half of the subjects manifested faster reactions to peripheral stimulation than to foveal stimulation. The other one-half, of course, reacted faster to foveal stimuli. The over-all group average, however, yielded a difference of 10 to 11 msec. in favor of the peripheral stimulus. Statistically, the indication was that 95 chances in 100 the difference was real. Figure 106 indicates the differences in reaction time under the four different conditions (two retinal and two stimulus).

There are a number of questions which we might ask in regard to reaction time under the two conditions. In what ways and to what extent were the findings dependent upon the levels of stimulation and upon the sense organs? It is easy to view this study as pertaining primarily to sense-organ or receptor behavior under two conditions of adaptation. It might be argued that the differences in the reaction times are brought about by the differences in the speeds at which the sense cells react. Some of the difference is undoubtedly brought about in this way. There is, however, the factor of readiness of the organism, as well as the rate at which sense organs react. Levels of general stimulation (nonspecific) have to do with readiness. That is, there is a difference in the over-all behavior of the individual in the dark and in the light. Devoid of the general excitation obtained in the light, the seeing individual attempting to detect suddenly appearing or weak visual targets reacts more slowly in the dark.

It is obvious that this study did not make any distinction between the sense-organ factor and the factor of lowered general excitation and added "uncertainty" in the dark. This might have been shown in some way by having the subjects react to nonvisual stimuli in both the light and the dark. If a time difference in favor of the reactions in lighted surrounds were to be shown, it would tend to substantiate the supposition that a material factor in the slower reactions in the dark is based upon a general

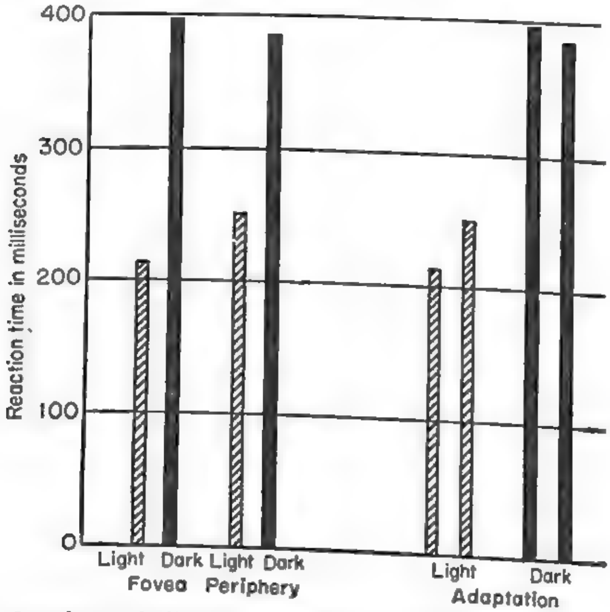


FIG. 106. Graphs to indicate reaction times under four different sets of conditions.

difference in the over-all state. This hypothesis, or another in this category, was not the central reason for Lemmon and Geisinger's study, hence we need not consider their failure to make the nonvisual tests an omission.

Reaction Time in Discrimination. Kellogg's study was one in which two halves of a circular disk were presented at equal or at unequal intensities. The disk subtended a visual angle of about 6 deg. and was viewed with only the left eye. To eliminate involuntary head movements, a chin rest and an eyepiece were provided. The judgments of the observer were made by pressing one of two or three keys, depending upon the number of

categories allowed. The judgments were recorded automatically on a kymograph drum by means of signal markers.

The apparatus was shielded from the observer's view, and only the telescope eyepiece protruded past the screen.

The subjects (observers) were five graduate students in psychology who each had made more than 4,000 psychological judgments before participating in this investigation. The subjects were ignorant of the purpose of the experiment including the fact that the times taken for responding were recorded. On this account, they were under no unusual pressure to respond quickly. This represents a marked difference from the usual reaction-time experiment. The absolute times taken to respond were not a crucial matter. It was the relative times taken to respond under different conditions that was of more particular interest.

There were seven pairs of intensity combinations, with values in arithmetic progression. In three, the left half of the disk was the more intense, and in three, the right half was the more intense. In one case the members of the pair were actually equal. The intensity was about 21.7 meter candles at equality at eye distance.

Judgments were made in series of 42. Between each series a 1-min. rest period was permitted. During this time the subject was asked regarding the judgments just completed. He was asked as to the percentage of the cases in which guessing played no part. So that undue emphasis was not placed upon one category at the expense of another, the subjects were instructed to look upon each category as being of equal importance. The subjects were instructed to act as though they were simply sorting cards into two or three boxes depending upon whether two or three sizes were involved.

Results. The intensity steps represented by the stimulus pairs are designated L_3, L_2, L_1, E, R_1, R_2, and R_3 in Fig. 107, which indicates the results. From the figure it is apparent that the longest times are required when there is the minimum difference in the objective intensities of the members of the pairs. In all, there was a total of 3,360 judgments, one-half of which were made on the basis of two categories and one-half on the basis of

three categories. Except for the judgments of one subject (in R_3), all judgments were quicker in the two-category trials than in the three. Thus the mere shift in attitude that requires the subject to be ready to judge between right and left, to one in which the subject must be ready to make one of three judgments, "right," "left," or "equal," considerably slows up the process of making a judgment.

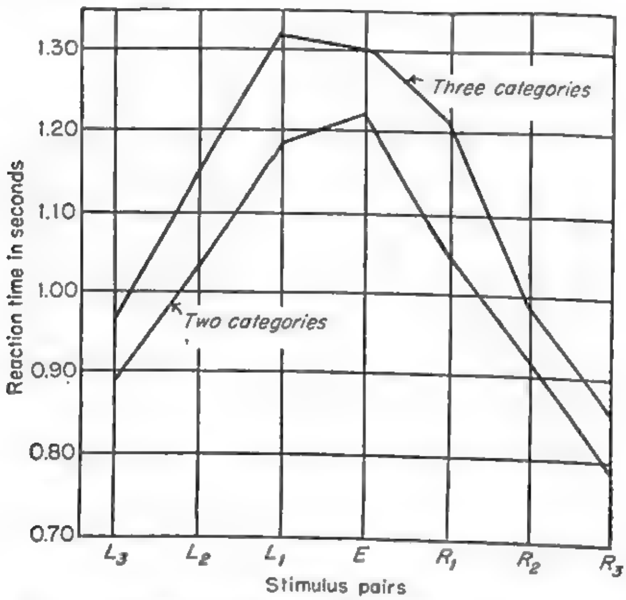


FIG. 107. Reaction times for various "constant stimuli" when two and three choices were allowed. (*After Kellogg. By permission of Amer. J. Psychol.*, 1931, **43**, 72, *Fig.* 1.)

It was apparent that the factors of (1) the kinds of choices possible, and (2) the extent of physical differences between the members of the stimulus pair were the major ones in determining the speed with which judgments could be made. Since speed was in no way mentioned, the differences are perfectly natural ones characteristic of performing the task.

Another very important finding was made. The records showed that it took longer on the average to make a false judgment than a correct one, regardless of whether two or three categories were involved. The differences between correct and incorrect judgments were found to be reliable, except for one subject

in his two-category trials. Even for him the difference was significant 996 times in 1,000.

Whereas the general tendency was to make quicker responses, the greater the physical difference between the members of the pair, this was not true for the cases in which incorrect judgments were made. Figure 108 indicates the findings when the judgments were separated into two kinds, "correct," and "incorrect." It will be seen that while the reaction times are shorter for the

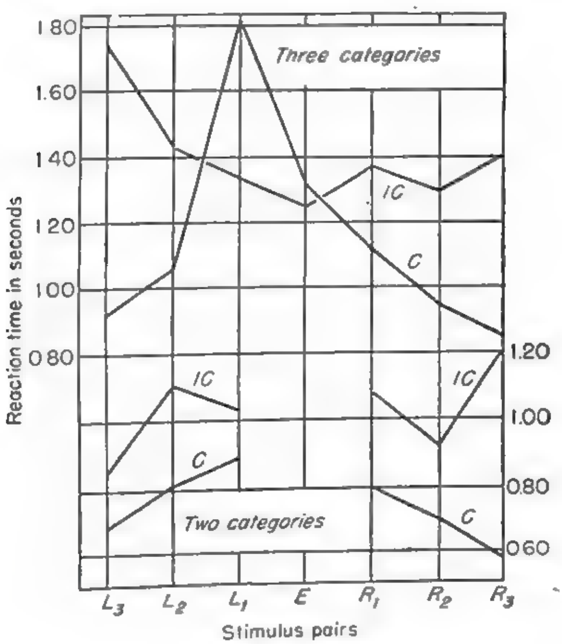


FIG. 108. A comparison of the reaction times for correct and incorrect judgments when Kellogg allowed two and three choices of response.

correct judgments as physical differences increase, this is not always true for the incorrect judgments. For the incorrect judgments for pairs R_1, R_2, and R_3 for both the three-category judgments, the results are irregular but show a tendency for the judgments to be longer for the greater physical differences.

It would have been instructive to have repeated the investigation with subjects (the same or different individuals) who were instructed to make their judgments as rapidly as possible, or as consistent with correctness or surety. Since the predominant

number of incorrect judgments took more time than the correct ones, would hurrying the subjects up increase the correctness of judgments, or would the whole level of accuracy drop? If it should drop, would there be a different relation in reaction times between correct and incorrect judgments?

It is obvious that the matter of certainty is involved in the process of making judgments such as in this investigation. Not all judgments made were made with equal certainty. Some closer tally on this factor might be worth while. The closest that the experimenters came to doing this was asking at the end of each series of 42 judgments regarding the percentage of guesses that were made.

The data showed that the greater average time in making equality judgments could not be ascribed to their being made most likely when the physical difference between the stimuli was small. The equality judgments that were made took longest when the difference between the stimuli was large. It was first suggested that the equality judgments, as a category of response, were longest because of the element of doubt involved, but reference to the subjects' reports indicates that the subjects felt no more doubtful in making equality judgments than either of the other kinds. An earlier study by George also bears out the likelihood that it is not the element of doubt that takes equality judgments so long to make. Kellogg finally concluded that it was neither the element of doubt nor the smallness of stimulus differences that made the reaction time so long for the equality judgments. It seemed that it lay in the very nature of making an equality judgment as in contrast to making those of inequality.

QUESTIONS

1. What is the object of presenting the studies given in this chapter?

2. What are the procedural essentials of a good reaction-time experiment?

3. What are the variables in reaction-time experiments?

4. What are the so-called "muscular" and "sensorial" reactions?

5. What two variables of "readiness" are mentioned in this chapter?

6. Describe the Lemmon-Geisinger experiment.

7. What additional variables might have been studied in that experiment?

8. Describe Kellogg's study.

9. What psychophysical method was employed in his study?

10. What measurement did Kellogg include that is not usual in such an experiment, and what were the findings obtained from its use?

CHAPTER 39

BLOCKING

Human activity has numerous characteristics. One of the most common characteristics of activity or performance is variability. Variability shows up to different degrees and in different ways depending upon the kind of activity in progress.

Some acts or responses are transient. They occur suddenly and are completed in a very short time—a matter of seconds. Other activity is prolonged. In the accomplishment of some tasks, activity may last for hours. Prolonged activity does not all belong in the same category. If motor, it may consist in sustained tensions or postures, or it may consist in an extended series of movements. These may consist either in repeated performances or in a series of quite different motions.

A series of acts or responses may be made in attempting to keep pace with some kind of a pacemaker, such as a metronome or a piece of machinery which the individual operates or feeds. In such cases the series of acts required of the individual is about as nearly repetitive as ever occurs. A modified form of so-called repetitive behavior is required in a series of choice reactions. For example, a series of colored squares of paper may be exposed at a given rate and the individual required to name the colors. There is always a decided limit to the number of colors used. Let us say, for example, that there are six colors. These are presented in random order, so that in a series of 48 presentations each color appears eight times. To this extent the presentations are repetitive. In the sense that all responses are names of colors, the responses are homogeneous. Even though the names of the six colors are to be repeated eight times each, it is possible for the individual to vary in his performance of pronouncing the

names and in the quickness with which he responds. Since choice reactions are involved in this case, some authors would call the performance mental.

One of the problems that arises in dealing with repetitive or quasi-repetitive response (a task sometimes called homogeneous) concerns the occasional failure to respond. In addition to the slight variation in rate of response as repetition continues, there are a few larger variations that could better be called pauses, or even outright failures. It is as if a kind of "block" had occurred in the underlying processes involved in the performance. In fact, such pauses have come to be called blocks.

In general, blocking refers to short periods during which individuals seem to be unable to perform. The manner in which blocking is manifested depends upon the nature of the task. In fact, whether blocking shows up at all or not depends upon the kind of performance involved as much as upon the individual himself. We shall define blocking in a more specific way later on, but let us consider now the possible significance of the phenomenon.

Bills, in investigating blocking, listed several significant considerations that call for a study of the phenomenon. (1) It was suggested by certain workers that failure of rapid reduction in achievement to occur in mental tasks, as in muscular work, was because the worker obtained frequent brief rest periods. These rest periods were presumed to give adequate chance for recovery. Recovery precluded accumulation of effects called fatigue or impairment and prevented a decrement in accomplishment. (2) It was long known that *attention* is characterized by a more or less rhythmic fluctuation. While this fluctuation has generally been dealt with as an experiential affair, Bills supposed that some corresponding fluctuation in overt performance would be found if only a sufficiently refined technique could be devised for the purpose. (3) Bills also mentioned the possibility of the refractoriness of nerve being at the bottom of periodicity in mental performance. The refractory phase to which he had reference was a possible accumulative refractoriness considerably greater in length than the exceedingly brief simple phase found in peripheral nerve fibers. (4) Bills also referred to the possibility

that blocking might be related to the appearance of performance errors in tasks calling for a series of responses.

It is our purpose, in this chapter, to deal with the investigation of blocking which Bills made.

After preliminary experiments, Bills chose five tasks that would seem to reveal the phenomenon of blocking in a fairly concrete way. He found that the tasks had to be those requiring a high degree of continuity and involving many brief responses, so that if blocking were to occur it would show up as gaps in the uniform flow of activity, or variations in the promptness of response. Blocking would be expected to cause the occasional omission of responses or their delay. Were the tasks to be composed of work dealt with in large units, the possible blocks would tend to be masked within the unit performances. The tasks chosen were (1) alternate addition and subtraction; (2) observation of reversible perspective; (3) color naming; (4) making substitutions; and (5) naming opposites.

Apparatus and Procedure. Fifty advanced students took part in the procedure. Since the procedure was essentially the same in all five performances tested, a single procedure description will be sufficient. The subjects were used one at a time. Directions were given to them on a typed card for uniformity. Each subject sat at a table with his back to the experimenter and apparatus. The subject responded orally to the stimulus material in front of him. He was urged to work as rapidly as he could, consistent with reasonable precision.

In front of the experimenter was the recording apparatus including the key which he pressed at the instant the subject made a response. A second person, acting also as a recorder, sat at the table and operated a second contact key to register every time the subject made an error. This individual was supplied with a correct-response list so that errors could be detected at once.

The two contact keys were connected to a kymographic recording system containing a continuously unrolling strip of paper. The keys electromagnetically operated markers of the fountain-pen type. Three markers were mounted side by side so as simultaneously to record responses, errors, and time in seconds

or in 3-sec. intervals. The timing marker was operated by a contact clock. The recording-paper speed was 6 in. per min.

For the purposes of the investigation, a block was operationally defined as "a pause in the responses equivalent to the time of two or more average responses." This mode of defining, rather than by a constant period, made it possible to use a criterion relative to the rate of each subject's performance. For example, an individual responding at the rate of 60 per minute would be manifesting a block were he to pause at least 2 sec. If his rate was only 30 per minute, a pause would have to be 4 sec. long to count as a block. Bills studied six aspects of the general problem of blocking, namely (1) the absolute length and frequency of blocks, and the relation of these two factors to the rate of responding; (2) the rhythmicity of blocking—the absence or presence of constancy in number and spacing per unit time. This included the tendency of response-bunching between blocks; (3) the rise or fall of frequency and length of blocks throughout the work period; (4) the effect of practice upon the frequency and length of blocks; (5) the relation of blocks and response errors; and (6) individual differences in response rate as related to frequency and length of blocks.

Addition-Subtraction. The task in this experiment consisted in the alternate addition and subtraction of the digit 3 for each in a long list of numbers (from 4 to 9) placed in continuous rows. There were 40 digits per row and 24 rows per page. One page, then, contained over 10 min. continuous occupation.

In this experiment, 10 subjects were used, each working 7 min. consecutively on each of 8 days. Errors were recorded for only two of the subjects. The average rate of addition and subtraction varied from 61 items in the first minute of work to 55.8 responses (or items) in the seventh minute. The number of blocks varied from 2.9 blocks in the first minute to 4.2 blocks in the seventh minute. The length of the blocks varied from 1.9 sec. in the first minute to 2.7 sec. in the seventh minute. Figure 109 indicates the effect of practice over an 8-day period upon the number of responses per minute, number of blocks per minute, and the average length of blocks in seconds. During the 8-day practice period an increase of about 100 per cent in

responses per minute occurred. During that time the number of blocks decreased about 20 per cent, and the average length of the blocks decreased by about 50 per cent.

Reversible Perspective. The second experiment utilized the phenomenon of reversible perspective. It was the subject's task to make the reversals occur as frequently as possible. The particular figure used was the classical "pile of blocks" that could

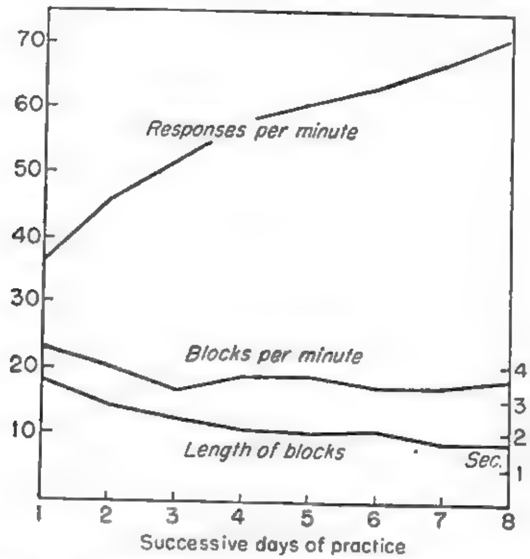


FIG. 109. Change in number of responses per minute, blocks per minute, and length of blocks with 8 days' practice in addition and subtraction.

be seen from above or below. The subject was presented with a card that contained the figure and a set of directions for procedure. In other respects the apparatus and arrangements were the same as for the previous experiment. Eight-day records were obtained for all of the subjects. Each record represented 7 min. work.

In the performance tested in the experiment, a 17 per cent decrease in frequency of response and a 35 per cent increase in frequency of blocks, during the 7-min. period, were found. When the average length of block for the first minute and the average for the last 3 min. were compared, a 7 per cent increase in length of block was found (see Fig. 110).

During the 8-day practice, the average increase in the number of responses per minute rose from 57.3 to 72 (about 26 per cent).

The blocks dropped from an average of 4.9 to 3.7 per min. (24 per cent). The length of the blocks remained almost the same (about 3 sec.). These averages are based upon 280 readings each.

Color Naming. Bills chose color naming as a task because it seemed to involve speed of apprehension, thus was taken to be a good test of perception. The colors of yellow, red, blue, green,

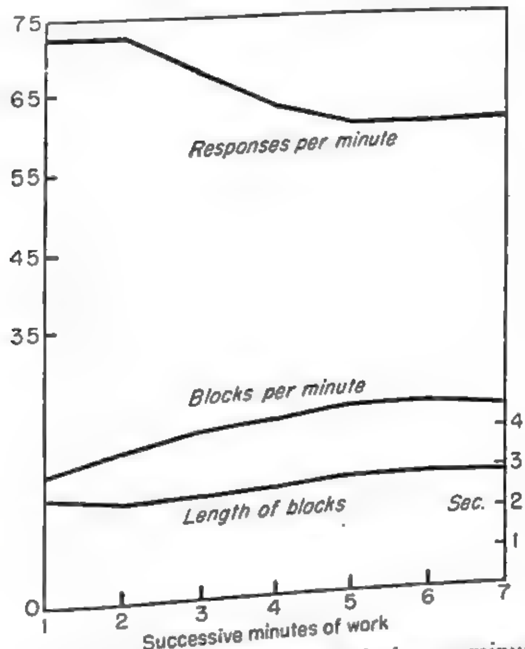


FIG. 110. Change in responses per minute, blocks per minute, and length of blocks with 7 successive minutes of looking at a reversible-perspective figure.

white, and black were exposed, one at a time, in all possible sequences.

For this experiment a new set of subjects was utilized. These were all uninformed regarding the purpose of the experiment. Each of the 21 subjects was given a task period of 10 min., on each of two occasions. The first period was used as a preliminary practice period, and the second period as the regular working period.

The results showed the same sort of tendencies as those of the preceding experiments. The number of blocks per minute,

during the 10 min., increased 33 per cent, and their length increased 18 per cent. The same rhythmicity of blocking showed up in this experiment as in the previous ones, and the same average interval between blocks (17 sec.) was again exhibited.

Substitution. The task in the present experiment was that of orally substituting letters for numbers in rows of random-order digits between 5 and 9. The substitutions to be made were, for example, *a* for 5; *b* for 6; *c* for 7, *d* for 8, and *e* for 9. The code was simple enough for easy mastery, but the task was repetitive or homogeneous enough for the supposed mechanism of refractoriness to operate. Errors were recorded in this experiment.

Twelve new subjects were chosen. Two days were used for practice, and 3 days on the test series. The work periods were 10 min. at highest possible speed. The average speed during the first minute on the third day was 92.6 responses. This fell to 83.3 for the tenth minute. The average number of blocks rose from 2 to 2.8, and the average length of blocks rose from 1.3 to 1.9 sec. In this experiment the average period between blocks was 24 sec. instead of 17 sec., as in the first three experiments. The stage of practice at which the records were taken may account for some of this difference. Rhythmicity in blocking appeared in this experiment just as in the previous ones. When individual records were treated statistically, a negative correlation between speed of responding and frequency of blocks was found. A negative correlation was also found between speed of responding and length of blocks.

The question of bunching of responses was given special attention in this experiment and the following one. Incidental observations seemed to indicate that between blocks there was a tendency for responses to bunch. By bunching is meant the tendency to be more frequent midway between blocks. This bunching or clustering, it was thought, might represent a rhythmic form of efficiency, the mid-points between blocks representing the crest of the waves and the blocks representing the troughs. The shapes of the waves also became a matter of inquiry. It was wondered whether the waves would be sinelike, or whether the waves would be abrupt in onset and the troughs broadly rounded.

For this purpose, certain records were taken with a faster paper speed so that the distances could be more accurately measured with a millimeter scale. Records were measured for the following purposes: (1) To determine the degree of uniformity in the recurrence of bunches and blocks; (2) to ascertain the shape of the waves; and (3) to determine the effect of practice and fatigue on the shape and regularity of the waves. Figures 111

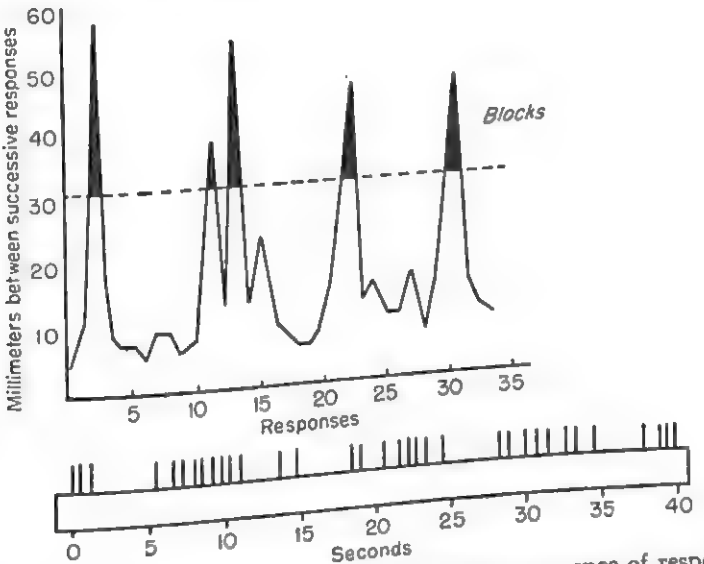


FIG. 111. Schema to show the relation of blocks to sequence of responses, before practice. The performance represented is number substitution. (*Bills. By permission of Amer. J. Psychol.*, 1931, **43**, 238, *Fig.* 1.)

and 112 give us a sample of the results obtained. At the bottom of each figure is a record indicating the temporal distribution of the responses. This information is also plotted in the form of a graph. The ordinate of the graph indicates the response sequence itself. Part way up the ordinate is a dotted horizontal line. This line represents the distance on the record beyond which the given subject could not pause without the pause being called a block. Hence all peaks in the graph rising above the dotted line indicate blocks. These peaks are blackened in so as to make the blocks more vividly discernible. Figure 111 indicates the performance of a given individual before practice, whereas Fig. 112 shows performance following practice.

Bills concludes that practice tends to make the peaks (blocks) occur less frequently, and that fatigue has just the opposite effect. Bills here, as elsewhere, in his blocking studies must mean by fatigue some sort of impairment that may or may not be recognized by the individual, but which is some fundamental result of mere activity.

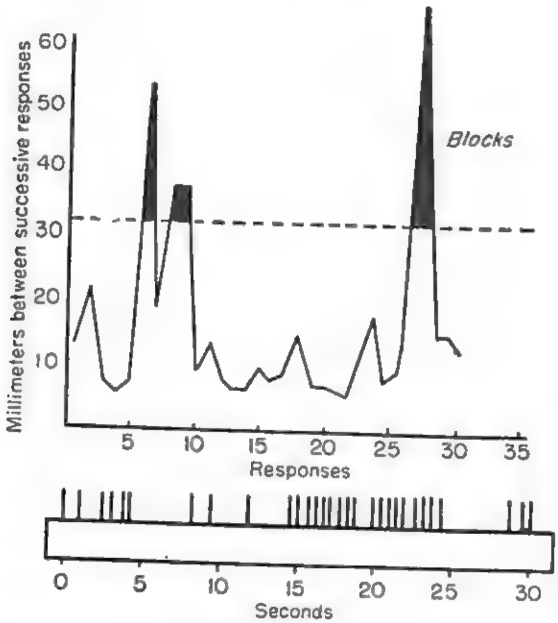


FIG. 112. Schema as in Fig. 111, but showing results following practice. In this case there was a greater bunching and fewer blocks. (*Bills. By permission of Amer. J. Psychol.*, 1931, **43**, 239, *Fig.* 2.)

Opposites. The results obtained in the experiment using naming of opposites as the task, turned out, in all essential particulars, to be identical with those in the previous experiment.

Conclusions. Bills, among other things, attempted to relate blocking to fatigue. Fatigue with which he dealt was not the subjective state that we should prefer to call fatigue, but rather some characteristic of the organism such as impairment of tissue or physiological function. While it is not certain what this so-called fatigue is, it is apparent that prolonged activity does result in certain changes in blocking. Blocking becomes more frequent, particularly when activity is continuously carried on for an hour. The duration of blocks also increases. The

frequency of response may, at the same time, remain almost uniform.

The concept of blocking and the technique of measuring blocks furnish us with a device whereby certain performances may be compared, and certain general effects on body state may be ascertained. This may be true even though we do not, as yet, have any good idea as to the mechanisms that underlie blocking.

QUESTIONS

1. Name and describe an important feature of variability of performance.
2. What kind of performances must be chosen to study blocking?
3. How did Bills define a "block"?
4. Describe Bills' apparatus and procedure.
5. What performance did he study?
6. What aspects of the general problem of blocking did Bills study?
7. Compare the findings obtained in the study of the several performances.
8. What do Figs. 111 and 112 show?
9. What relation did Bills conclude existed between blocking and fatigue?
10. What relation did blocking have to learning?

CHAPTER 40

AN INVESTIGATION ON FATIGUE

The history of the concept of fatigue covers an indeterminable length of time. No one knows how far back man came to distinguish between his feelings to such an extent as to set aside a definite category, somehow related to effort, which we could identify as representing our present category of fatigue. It is certain, however, that fatigue is the man-in-the-street's invention. It has become a laboratory category only by adoption.

In this adoption it has become greatly modified. It is now one thing to one group; another thing to still other groups. It has been given a number of widely different definitions. What is disappointing is the fact that most effort given to the study of fatigue has resulted in the inclusion of an ever-broadening heterogeneity of more or less unrelated phenomena all labeled by the one word—fatigue. Little headway has been made in clearing up the many opposing and incompatible features involved in the phenomena thus assumed.

An analysis of the thinking in this area over the past fifty or more years would provide an enlightening commentary on the progress of many aspects of biological sciences of modern times.

We are presenting as an example of fatigue study, one in which, to begin with, all of the features, good or bad, of conventional theory regarding fatigue were inherent. As the work progressed, the negative results were impressive enough to induce the investigators to do some individual thinking. As a consequence, the conclusions reached were very much different from virtually all those that have been made anywhere before or since.

The work is then an example not of proving a hypothesis but of demolishing certain aspects of it and paving the way for a totally different kind of thinking.

The investigation for our attention is that of Whiting and English. In part I of their study, six tests were used. Three of them were primarily motor, and three were "mental." As a group, the tests were devised to ascertain accuracy, difficulty, and speed of performance. The subjects were 16 young women in an Eastern girl's college. The tests were given at 8:30 A.M. or before, and again at 4:30 P.M. All subjects were tested on from 3 to 5 days. The tests given in the morning and in the afternoon were equated as nearly as possibly in difficulty, when they were not actually the same. The investigators used the same presuppositions regarding fatigue and the day's activities, as were given in the example in Chap. 3, which describes the logical basis of experimentation. Whiting and English assumed that during the day the subjects became fatigued by their ordinary routine activities, and that differences between afternoon and morning performances would indicate this.

Part I. *Test One.* In this experiment, the attempt was made to test the effect of fatigue upon the accuracy of arm movement. The blindfolded subject ran her pencil along a straight edge for a fixed length determined by a stop. It was the subject's task to reproduce as accurately as possible the length of line traced in the first motion. The exact length of the movement was recorded graphically. The lines used as standards were 7.4, 10, and 12.6 cm. in length.

The results indicated very little difference in performance between morning and afternoon. In 50 out of 90 cases, the reproduced motion was *shorter* in the morning than in the afternoon. On the other hand, the reproduced motion was longer than the *standard* in 74 cases in the morning and in only 61 cases in the afternoon. The average error in the afternoon trials was 9.6 mm., and in the morning it was 9.3 mm.—actually, not a significant difference.

Test Two. This test was supposed to ascertain the effect of fatigue on mental accuracy. The subject was required to bisect and trisect four lines (8.3, 15.3, 14.2, and 9.0 cm.) by visual inspection. Much the same results occurred in this experiment as in the previous one. There was no evidence to indicate that accuracy diminished during the day.

Test Three. This experiment was intended to test the effect of fatigue on the speed of intellectual activity. For this, addition tests were used. A column of 3's and 4's in random order was given to the subject. These digits were cumulatively added, one at a time, to a number given as a starting point different for each test. The process of addition was continued for 30 min., with the instructions to work as fast as possible.

Despite emphasis on speed, very few errors were made. In 31 cases out of 108, there was less accuracy in the afternoon; in 17 cases there was greater accuracy; and in 60 cases there was equal accuracy. There was a negative relation between reports of being fatigued and the degree of accuracy in the performance. The speed was greater in the afternoon than in the morning. In the morning there were 19.5 items added, per trial, as against 21.0 in the afternoon.

Test Four. This experiment was meant to test the effect of fatigue on the difficulty of performing a physical task. The subject was required to hold a stylus at arm's length in each of three holes, in a Whipple steadiness tester, consecutively. The diameters of the holes were $^{11}/_{64}$, $^{10}/_{64}$, and $^{9}/_{64}$ in. Inability was indicated by unavoidably touching the sides of the holes. In every case this completed an electric circuit and was recorded on a kymograph. Both arms (hands) were tested. The subjects did a little better in the afternoon than in the morning. In 85 cases, fewer contacts were made in the afternoon. In 57 cases, more contacts were made; and in 26 cases, the same number of contacts were made.

Test Five. This experiment was designed to determine the effect of fatigue on the difficulty of an intellectual task. Ten problems in multiplication, progressively increasing in difficulty, were presented, one at a time, through a small aperture in a screen. The subject was required to solve them without the use of pencil and paper. It was found necessary to limit the time required to do the problems to 15 min. All but one subject generally finished within the required time. In this test, the accuracy was practically the same in the afternoon as in the morning.

Test Six. In this test, the subject was required to tap as rapidly as she could upon a metal plate with a stylus, each tap being recorded electromagnetically. A slight but insignificantly

greater number of taps per unit time were made in the afternoon.

In all of the tests reduced performance failed to occur in the afternoon. Thus, although it was firmly held that fatigue had occurred during the ordinary activities of the day between 8:00 A.M. and 4:30 P.M., no evidence of this showed up in the tests. Such a discrepancy caused the investigators to ask both themselves and their subjects whether the assumption that fatigue developed during the day was wrong. The subjects did not think so.

The subjects definitely felt tired during the afternoon test-period. There was actually considerable discrepancy between the way they felt and the level of their test-performance. Sometimes the tired subjects performed better than they did when they were not tired. We must realize that the subjects knew they were participating in a fatigue investigation and were thus able to throw themselves into their activity with added zest to offset the way they felt.

Since all the results were negative, in a sense, the investigators decided that further experimentation with modified procedures was called for. They felt that one of the major things that was needed was a means of offsetting this tendency to spurt and therefore possibly mask the effects of fatigue.

Part II. *Further Tests.* Accordingly, a second set of experiments was planned. These experiments involved prolonging the test-activity period to offset the effect of possible spurts.

A series of performances were devised, most of which were much like or identical with the tests in part I. These were planned to keep the subjects busy for 90 min. No opportunity was allowed the subjects to recover from any extra effort put forth in any single test. The test-period was placed in the afternoon. The tests covered a variety of kinds of performance including reproduction of lines, demonstration of steadiness, memorizing nonsense syllables, demonstration of speed of movement, cancellation, etc. The battery was planned to require 45 min., at the end of which the subjects went through the battery in reverse order. During the 90 min., each kind of performance was repeated several times, so as to provide information as to any possible decrement in performance of a given kind.

The broad outcome of this new experiment was about the same as the outcome in the six tests of part I. Since no decrement was actually found, the plan of the investigators to test the effect of incentives on fatigue could not be carried out. Incentives, of course, had already been involved. These consisted in giving the subjects knowledge of results after much of the experiment had been performed. In the final repetition of the test-items, some effect of this "incentive" was evident in the form of greater speed. But less accuracy went with the increase in speed. Accordingly, the net result was not one of gain. Thus there was no evidence of incentives having improved performance.

Summary and Conclusions. We may now say more about the effect of the investigation upon the investigators, adding some comments of our own.

Despite the negative results in both parts I and II, the authors could not believe that the absence of evidence for fatigue was because there was no fatigue. The authors believed that there was no plausible meaning that could be given to the word fatigue that would exclude the development of fatigue during the day's work. In fact it was believed that performance in the experimental situation would also involve fatigue, thus adding more fatigue to that which was assumed to have already developed during the day.

While the outcome of the investigation was not an array of quantitative material giving evidence toward a "proof" of some hypothesis, it led the investigators to do some more thinking upon the matter. In this respect it was a very good outcome, possibly better than if a scant amount of data had been used to bolster some specific idea.

Finally, Whiting and English concluded that fatigue is a subjective affair, and as such must be akin to the appetites. It is a complex-feeling state involving emotional coloring, visceral and organic sensations, etc. The authors labeled fatigue, finally, as an emotional appetite. Such a category must have seemed rather unusual at the time the authors published it.

Whiting and English did not go ahead and describe the essential characteristics of appetite so as to draw a detailed parallel between them and the characteristics of fatigue. It might have

been easily possible to have done so. With some exposition on the subject, a quite convincing case might have been made for fatigue as a "negative" appetite. But lacking that demonstration, we should tend toward putting fatigue in a class such as that of the attitudes. It is to be recognized that even the term "attitude" is somewhat unsatisfactory, due partly to the number of ways the term has been used. It has been used not only to signify the involuntary and pervasive orientation of the individual emotionally, etc., but also to signify the temporary and willful positions he takes upon occasion for specific reasons. Since, according to convention, this lighter meaning is common, the word has lost its significance as a label for the characteristic of orientation involving the whole organism. To make amends for this varied use of the term, lately the term "stance" has been used by one or two authors. They wish the term stance to have none of the superficiality now suggested by the term attitude, but to signify the alignment of the entire organism.

Fatigue, when classed as a stance, represents the individual's self-appraisal with regard to his own abilities, comfort, and dislike for certain activities. Even so, it is not an intellectualization but a condition arrived at as an expression of the relation between the factors just mentioned. It includes a bodily expression of this relative revulsion and inadequacy, due to the fact that the stance, in the first place, pertains to activity. Since activity is carried on by physiological mechanisms, the aversion and inadequacy is expressed through tangible changes in these mechanisms.

Whiting and English pointed out what to them, at the time, seemed to be a fundamental distinction. They believed that a distinction between *fatigue* and what they called *exhaustion* should be recognized. Whereas exhaustion is basically a physiological or frequently a localized phenomenon, fatigue is to be taken as a psychological one. Partial exhaustion always results in reduced capacity for work. Whereas exhaustion pertains to the mechanism whereby work is accomplished, fatigue has to do with the drive for work. We may add that a lack of a one-to-one relation between the two might well be expected, and that quite dissimilar experimental procedures should be used to disclose the presence or absence of the two.

In making this distinction it can be said that Whiting and English made a very substantial contribution to the thinking regarding the fatigue problem. The value of the tests which they made lay in showing them that their initial concepts did not fit reality. These authors were able to take advantage of their experimental findings. This was in sharp contrast to the usual outcome, for general opinion about fatigue has not yet reached the point where fatigue and impairment are systematically distinguished.

This investigation is an example of how a set of experiments will enable an experimenter or a small team of workers to arrive at a conclusion that for some reason or other does not impress those who read the report in the literature, despite the fact that the conclusions open up a whole new field of possibilities. The distinction between fatigue and impairment (exhaustion) was a perfectly logical one to make, whether it came solely from the investigation Whiting and English made or from other considerations. It is possible that the authors did not, at the time, see so clearly the significance of the distinction they made as is now possible, else they might have pursued the matter further by having more to say or by performing more experiments.

QUESTIONS

1. What were some of the assumptions that underlay Whiting's and English's study of fatigue?
2. Name the tests Whiting and English performed in Part I of their investigation.
3. Were the performances of the subjects better or poorer in the afternoon than in the morning, or was little difference shown?
4. Why did the investigators include a second part to their investigation?
5. What were the tests used in Part II?
6. What were the results obtained in this part?
7. What did Whiting and English conclude regarding fatigue?
8. How did their conclusions differ from the general conclusions made from fatigue experiments before and since that time?
9. What contrast did the investigators make between fatigue and exhaustion?
10. What was Whiting's and English's main contribution to the thinking in regard to fatigue?

PART IX

SOCIAL PSYCHOLOGY

CHAPTER 41

THE SOCIAL BASIS OF PERCEPTION

It is coming to be more clearly recognized that we need to know considerably more about perception. We need to know how it is dependent upon other processes and in turn how these processes are dependent upon perception. As we give attention to these two aspects of this very fundamental problem, we must avoid thinking of perception as a detached or isolated process. Perception is the whole individual in action. It is the name for the active process by which the individual makes contact with and immediately responds to the external physical world. For the individual, the physical world is largely a social world. Hence to study perception is to study how he behaves socially.

Earlier concepts of perception made it a kind of embellishment or modification of a process called sensation. Just what sensation was could never quite be defined in a satisfactory way. It was supposed that arousal of sense organs led to an afferent input into the central nervous system that, under special conditions, could end up as an experience devoid of meanings and connections with past experience, etc. Try as men would, the theorizers and experimenters never actually succeeded in producing such experiences. Gradually arbitrary distinctions between sensation and perception have fallen into disuse. Few if any such hypothetical distinctions find much support in conviction these days.

During the time that considerable effort was being expended in attempting to isolate sensations as structural units of mind, a growing number of workers gave their attention to matters that seemed more plausibly to concern personality structure and human behavior. The emphasis in psychology left the realm of sensation, thought, and attention and shifted to matters of emotion, wishes, etc. Although the desire to know about emotions,

wishes, and conflicts was both natural and laudable, it left the pressing problems regarding the organisms' contact with the physical world largely unsolved. It seemed to many as though the world in which man lived was, for all practical purposes, social and not physical. This being the case, there seemed to be more direct ways of studying social phenomena. All of such thinking overlooked the fundamental nature of perception in relation to social behavior.

More recently a revival of interest in perception has begun to show itself. Perception is slowly coming to be seen as underlying all other functions, and in turn is an expression of these functions. Now perceptions are not to be distinguished from sensations. Since we can find no distinguishing characteristics, we have to say that sensations (if we insist upon keeping the term) are modifiable. The modifiability of sensation by training or experience is the recent contribution to sense-perception theory.

It is the purpose of the present chapter to describe an investigation in which "value" and "need" play roles in organizing perception. Since both value and need have social aspects, the investigation to be described is thought of as being a definite contribution to social psychology.

As an introduction to the investigation of Bruner and Goodman, we shall reiterate what they have to say regarding the determinants of perception. These authors wish to distinguish between two types of determinants. The one they call *autochthonous* (indigenous or native); the other, *behavioral.* Under the former label they group the properties of the nervous system that make for highly repeatable and, therefore, predictable response. Among the phenomena determined in this way are flicker fusion, binaural beats, paradoxical warmth and cold. The authors point out that given ideal darkroom conditions the average organism responds to a set of conditions in relatively fixed ways. We might add that it is the relative stability of all the conditions involved that makes for a relatively fixed outcome. Sets of conditions which one can set up in the laboratory or which one meets in everyday life vary all the way from this stable pattern to unimaginable extremes. It was stable conditions which led to fixed results that fooled those who used such

conditions in their sensory investigations. The false notion was gained by such investigators that since the results were stable they were dealing with an intrinsically different order of phenomena. Consequently, they had to have a term to label and distinguish such phenomena from the more fluid and labile conditions encountered elsewhere. This gave birth to the concept of *sensation* and the distinction between sensation and perception.

Bruner and Goodman put into the category of behavioral determinants of perception such processes as the adaptive functions of the organism (the learning processes, the dynamics of repression, the social needs and attitudes, and so on). Those studying processes such as conditioning have demonstrated that individuals can be conditioned to *see* and *hear* things in about the same way that they can be conditioned to perform overt acts, such as bodily movement (eye-blinks, knee-jerks, etc.) (Ellson and others). For example, if a sound and a faint visual stimulus are paired, the visual object will be seen later on when omitted from the pair. Whether or not this be called suggestion by the old terminology is beside the point. The subject has been conditioned so as to respond to conditions other than those originally necessary for the perception of an object.

Other experiments have demonstrated that the *same object* may come to be perceived *differently* (Proshansky and Murphy, and others).

Lastly, Bruner and Goodman bring out the fact that the world in which the organism lives is filled with more or less ambiguously organized stimulus material. What the organism perceives, in any case, is the result of some sort of a compromise between autochthonous and behavior processes. This is to say that, in many cases, the physical conditions are such that no violence is done in the "interpretation" of physical externality, regardless of whether a perception has one value or another. For example, an illuminated rectangle in a darkroom may be perceived as a small object close by or as a much larger one some distance away. No violation of optical laws is involved, regardless of the perceived size of the rectangle. Nature contains many ambiguous situations, situations in which the organism is left to use any one of a number of alternatives of its own.

Hypotheses. Bruner and Goodman made several hypotheses which they decided to put to test. Their three hypotheses are the following.

1. The greater the social value of an object, the more it will be subject to organization by behavioral determinants. The object will be selected from among other perceptual objects. It will become fixated as a perceptual-response tendency. It will tend to become accentuated in some way as a perception.

2. The greater the individual need for a socially valued object, the more accentuated will be the operation of behavioral determinants.

3. Perceptual equivocality will aid or enhance the operation of behavioral determinants, but only to the extent that it reduces the operation of autochthonous determinants. Perceptual equivocality is expressed, for example, in behavior toward certain geometrical figures. These we call ambiguous figures—figures with reversible perspective. Evidently, as will be seen later, the authors consider response to physically absent objects as involving the matter of equivocality.

The Experimental Conditions. Bruner and Goodman used 30 ten-year-old children, normal in intelligence, as subjects. These were divided into two experimental groups and one control group. The apparatus was a rectangular wooden box (9 by 9 by 18 in.). At one end there was a ground-glass screen 5 in. square. At the same end was a knob at the lower right-hand corner. Turning this knob varied the size of a special bright disk of light (16.2 foot-lamberts) from ⅛ to 2 in.

The subjects were used individually. The box rested on a table, and the disk was slightly below eye level. The experimenter sat behind the box. He told the subject that this was a game and that the subject was to make the light disk on the box the same size as various objects shown or told about. Before beginning the experiment proper, the subject was permitted to see how large and how small the disk could be made.

In both experimental groups (20 children, total) the subject was first asked to estimate the sizes of coins from a penny to a half dollar. This was to be done by adjusting the disk to the sizes that would represent their judgments. The series of coins

followed the ascending order of value, then the descending order. Two judgments were made for each coin, one beginning with the disk at its minimum size; the other beginning with the disk at maximum. Thus, four judgments were made by each child for each coin. In every case this was without the subject's subsequent knowledge of how close his judgment had been to the actual size of the coin.

The second procedure was the same as the first, except that the coins were now present. Each coin, one at a time, was held in the palm of the left hand at a level with the light disk, and 6 in. away. Each subject was allowed as much time as he desired.

A control group of 10 subjects followed the same procedure as just described, but instead of coins, disks of medium-gray cardboard were substituted. The sizes of these disks were, of course, identical to those of the coins. No mention of money was made in the presence of these individuals.

Results. The first comparison to be made was between the results in judging cardboard disks and coins of the same size. Figure 113 (left) indicates the outcome. It will be seen that a very different estimation was made for the two. In the case of

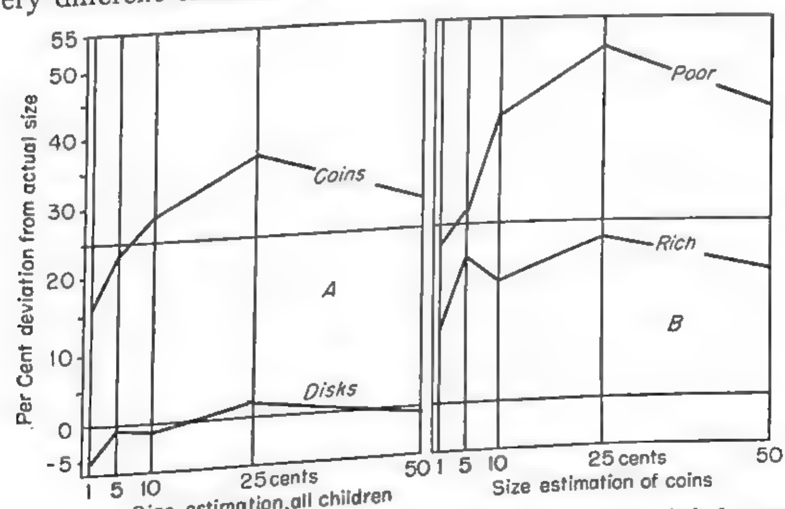


FIG. 113. Judgments of sizes of coins and disks by children (whole group, left) and judgments of coins by poor and by well-to-do children within group (right). (*Bruner and Goodman. By permission of J. abnorm. soc. Psychol. and Amer. Psychol. Assn.*)

Hypotheses. Bruner and Goodman made several hypotheses which they decided to put to test. Their three hypotheses are the following.

1. The greater the social value of an object, the more it will be subject to organization by behavioral determinants. The object will be selected from among other perceptual objects. It will become fixated as a perceptual-response tendency. It will tend to become accentuated in some way as a perception.

2. The greater the individual need for a socially valued object, the more accentuated will be the operation of behavioral determinants.

3. Perceptual equivocality will aid or enhance the operation of behavioral determinants, but only to the extent that it reduces the operation of autochthonous determinants. Perceptual equivocality is expressed, for example, in behavior toward certain geometrical figures. These we call ambiguous figures—figures with reversible perspective. Evidently, as will be seen later, the authors consider response to physically absent objects as involving the matter of equivocality.

The Experimental Conditions. Bruner and Goodman used 30 ten-year-old children, normal in intelligence, as subjects. These were divided into two experimental groups and one control group. The apparatus was a rectangular wooden box (9 by 9 by 18 in.). At one end there was a ground-glass screen 5 in. square. At the same end was a knob at the lower right-hand corner. Turning this knob varied the size of a special bright disk of light (16.2 foot-lamberts) from ⅛ to 2 in.

The subjects were used individually. The box rested on a table, and the disk was slightly below eye level. The experimenter sat behind the box. He told the subject that this was a game and that the subject was to make the light disk on the box the same size as various objects shown or told about. Before beginning the experiment proper, the subject was permitted to see how large and how small the disk could be made.

In both experimental groups (20 children, total) the subject was first asked to estimate the sizes of coins from a penny to a half dollar. This was to be done by adjusting the disk to the sizes that would represent their judgments. The series of coins

followed the ascending order of value, then the descending order. Two judgments were made for each coin, one beginning with the disk at its minimum size; the other beginning with the disk at maximum. Thus, four judgments were made by each child for each coin. In every case this was without the subject's subsequent knowledge of how close his judgment had been to the actual size of the coin.

The second procedure was the same as the first, except that the coins were now present. Each coin, one at a time, was held in the palm of the left hand at a level with the light disk, and 6 in. away. Each subject was allowed as much time as he desired.

A control group of 10 subjects followed the same procedure as just described, but instead of coins, disks of medium-gray cardboard were substituted. The sizes of these disks were, of course, identical to those of the coins. No mention of money was made in the presence of these individuals.

Results. The first comparison to be made was between the results in judging cardboard disks and coins of the same size. Figure 113 (left) indicates the outcome. It will be seen that a very different estimation was made for the two. In the case of

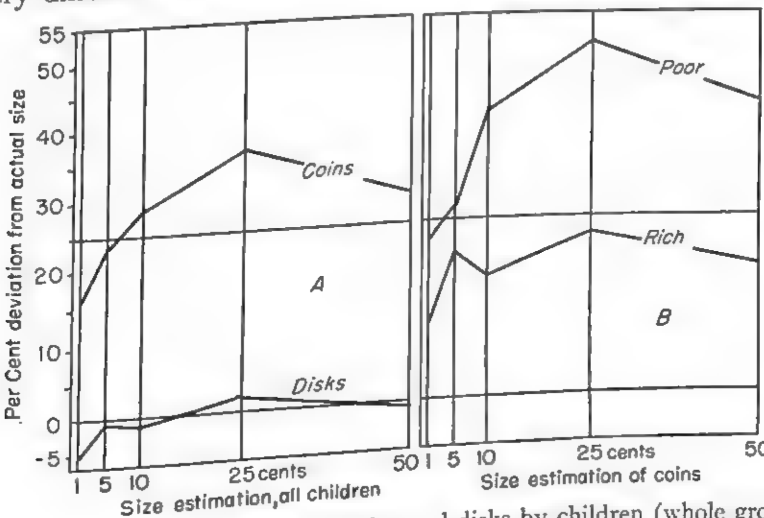


FIG. 113. Judgments of sizes of coins and disks by children (whole group, left) and judgments of coins by poor and by well-to-do children within group (right). (*Bruner and Goodman. By permission of J. abnorm. soc. Psychol. and Amer. Psychol. Assn.*)

the cardboard disks, the judged sizes were not very different from the actual sizes of the objects in the series. The coins, however, were all judged considerably larger than they actually were. The overestimation increased as the value rose from a penny to a quarter. For the half dollar, a reversal in the trend occurred. While it is not known why this reversal occurred, the authors hazarded the guess that the half dollar had a diminished reality value. They also suggested the extreme likelihood of the operation of some autochthonous reason.

The authors subjected their findings to statistical treatment to make sure of the validity of the sampling represented by the judgments. The treatment assured them that the results were to be relied upon. To explain the statistical procedures actually used is beyond the province of this chapter. You will be introduced to such matters in your course in statistics.

We may say that the findings support the first hypothesis that Bruner and Goodman made, namely, that the perception of socially valued objects is dependent upon behavioral determinants in proportion to their value.

The second hypothesis, that the greater the personal need for a socially valued object the greater will behavioral determinants influence perception, was studied by dividing the experimental group into two subgroups. One was called a rich group, and the other a poor group. These were equal in number; each contained 10 children. Subjects from well-to-do families were obtained from a private school catering to children of prosperous professional and business people. The poor children came from a settlement house in the city slum area. The investigators assumed that the poor subjects would have a greater subjective need for money than the rich children. Figure 113 (right) indicates the results of dividing the coin scores represented in Fig. 113 (left) into two groups. The poor children overestimated the size of the coins much more than the rich children did. The authors were unable to explain the dip in the rich children's curve for the dime. Curves plotted for more than two thousand adults also contain the dip. The investigators suggest that it might be due to the discrepancy between the comparative size and value of the dime, or that it could be due to something inher-

ent in the appearance of the coin itself. A randomizing of the order of presentation might change the results somewhat.

Statistical treatment of the results for the comparison of the rich and the poor children's choices also assured considerable reliability in the sampling.

The investigators assumed that they were testing their third hypothesis by comparing the memory estimations of coin sizes and judgments made with the coins present. That they were

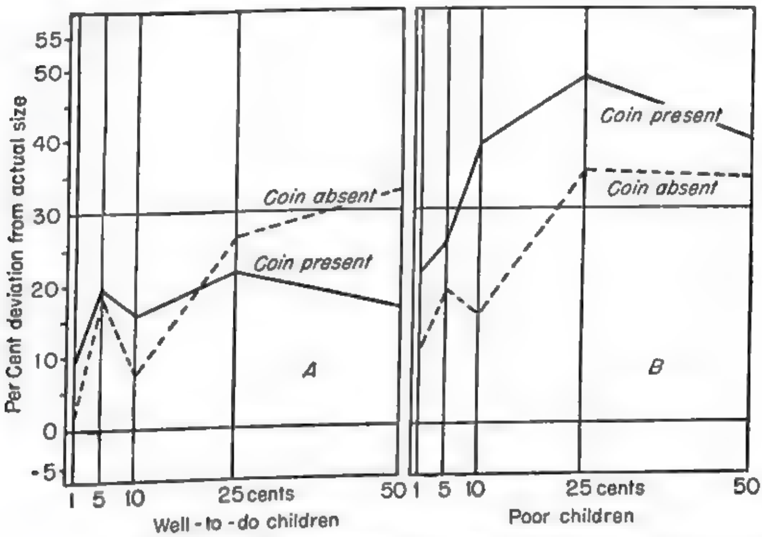


FIG. 114. Comparisons of judgments of sizes of coins when present and absent, by poor children (right) and well-to-do children (left). (*Bruner and Goodman. By permission of J. abnorm. soc. Psychol. and Amer. Psychol. Assn.*)

testing the matter of equivocality as implied by the hypothesis may be questioned. Nevertheless, we can examine the results obtained. Figure 114 (left and right) indicates the size estimation, with coins present and absent, for the rich children and for the poor children respectively. The dip for the dime again shows up in most of the curves and is not dependent upon the presence of the dime as a reference. Further research is obviously necessary to explain the outcome from comparing the coin-size estimation from memory and from observation.

In conclusion, it can be said that in the present investigation we have represented one of the many features of perception that

needed to be explored. The present study was a very good beginning and confirmed the few other more or less recent studies on perception that fixed universal results are not to be expected, even in situations in which the task is the standard routine of judging physical size of objects. The organism is not a neutral meter stick but a system that responds to the physical world as a world of value. The values are those largely determined by the individual himself.

QUESTIONS

1. Compare earlier views with the present outlook with regard to sensation and perception.
2. Are sensations unmodifiable reactions?
3. What is the purpose of the present chapter?
4. What types of determinants of perception did Bruner and Goodman distinguish?
5. What were the hypotheses that these investigators tested?
6. How did they go about testing them?
7. What kinds of subjects were used?
8. What were the findings of the Bruner-Goodman experiments?
9. What significance would you attach to these findings?
10. Design another experiment that might be expected to show the same principles.

CHAPTER 42

A COMPARISON OF SMALL GROUPS AND INDIVIDUALS IN CERTAIN PROBLEM SOLUTIONS

One major problem in social psychology pertains to the comparison of groups and individuals in the accomplishment of intellectual tasks. In the performance of routine tasks it is more nearly obvious that several trained individuals will accomplish a certain bulk of work in less time than required by a single individual. The question remains, however, as to what several individuals can do when simultaneously working in an intellectual problem-solving situation in comparison to a lone individual involved in the same task. There are numerous kinds of tasks, and it is possible that a comparison between individuals and groups would not give similar results for each one of them. For example, there are problems that have more than one solution, and there are others for which there is only one right answer. These two kinds of problems might result in different outcomes. We have at hand, however, a comparison between small groups and individuals in the rational solution of a complex problem having only one right answer. Such a study was made by Shaw in 1932.

Shaw's Study. Shaw presented to individuals and groups, containing four people each, problem situations calling for real thinking to accomplish their solution. The problems she chose all involved a series of steps each of which had to be correct in order that the right answer might be reached. She divided her investigation into two halves. In each half there were three problems, and each of these problems was given to a group and to four individuals. The subjects were psychology students in a university. Most of them were graduate students working for

advanced degrees. For the first half of the investigation, there were two groups containing 4 women each, and three groups of 4 men each. Working as individuals, there were 9 men and 12 women. In the second half of the investigation there were two groups of 4 men, and three groups of 4 women, each. There were also 10 men and 7 women working individually. A group never contained both men and women, for it was believed that groups of one gender would make for better cooperation and smoother operation. No sex comparisons were intended by this segregation and none were made.

First Half of Investigation (**Problems**). *Problem 1.* The group was given an envelope in which there were six disks labeled H_1, H_2, H_3, and W_1, W_2, W_3, to be used to represent husbands and wives respectively. On a card the problem was stated. It ran something like this: On the one side of a river were three wives (W_1, W_2, W_3) and their husbands (H_1, H_2, H_3). Each one of the men and none of the women could row. Your task is to get them across to the other side by means of a boat carrying only three persons at a time. No man allowed his wife to be in the presence of another man unless he was also there.

Problem 2. This problem was somewhat similar to problem 1, but concerned missionaries and cannibals.

Problem 3. The materials for this problem were in another properly labeled envelope. The problem had to do with arranging a series of disks in order of size by transferring them from one location to another so that the disks in the new location would have the same order as in the old. The old location consisted in a circle A. A second circle B was used as a transfer station, and the final location of the disks was to be in circle C.

Second Half of Investigation (**Problems**). Two weeks after the first, the second half of the investigation was undertaken. It consisted, as was said, in three problems. Each of these was different from those in the first, and different from each other. The card stating the first problem (1) said that in New Orleans there was a tree that nobody ever saw without becoming curious and wondering how it came to be there. The tree reminded people of the warm climates of Asia and Africa. It had sharp and thin foliage. Under the blast of northern November winds

it seemed to sigh mournfully and thus looked as sorrowful as an exile. In the envelope for this problem was a group of words that had to be put together properly to form the last sentence of the unfinished prose selection just described.

Problem 2. The materials were in a labeled envelope. The problem consisted in using the given group of words to finish out the three and one-half lines of an unfinished sonnet.

Problem 3. This problem was one of finding the most desirable location for a rural consolidated school within an area containing a complex network of roads. The mileage between each intersection was indicated on a diagram, as shown in Fig. 115. The

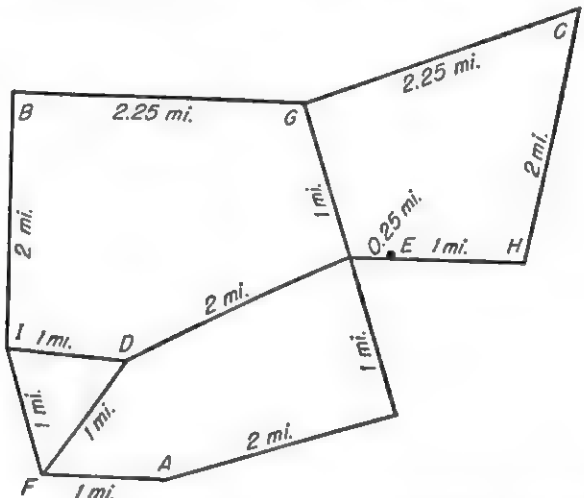


Fig. 115. Diagram of school district. (*From Shaw. By permission of Amer. J. Psychol.*, 1932, 44, 494, *Fig.* 1.)

capacity of each of two school busses was indicated, and it was stated that the busses might start from any point in the territory and did not have to start from the same point each morning. The number of children to be picked up at each intersection was also indicated. In solving the problem of finding the most desirable location for the school, the route each bus must take had to be indicated by the subject or group.

The general instructions that applied to all of the problems indicated that they were to be worked as quickly and as accurately as possible. The instructions also stated that there was a correct solution to each problem. In solving the problem, the

solver was to state briefly how the solution was made. When the first problem was finished, the second was immediately to be attempted. The solvers were to record the time to the nearest ½ min. from the record being kept on the front blackboard. Additional instructions were given to the *groups* doing the solving. They were approximately as follows: One member of your group has been appointed as chairman and is to manipulate the necessary materials. You are to operate as a cooperative group to solve the three problems that you have been given. Work them as quickly and as accurately as you can. There is a right answer to each problem. When you record your solution, state precisely how you solved the problem. Each person including the chairman should offer his contributions to the group spontaneously as soon as they occur to him. Indicate to the note-taker (an additional person who did not participate in any way in solving the problem) when you have completed one problem and are ready to go to the next.

An additional precaution in experimental design involved the equalization of the groups and individual workers in ability. The attempt was made to be reasonably sure that no group was composed of four superior individuals and others of inferior ones. It was also necessary to see to it that the individuals who worked in groups were, in general, neither inferior nor superior to those working separately.

Results. Table 19 indicates the number of individuals working the three problems in the first half of the investigation. It also indicates the number of groups working on the problems. Since there were 21 individuals and 3 problems, the total number of solutions at best could have been 63. As it was, however, only 5 correct solutions were made. In 50 instances, incorrect solutions were offered; and in 8 instances, the problems remained unsolved. Since there were 5 groups and 3 problems, the total number of solutions at best could have been 15. Actually there were 8 solutions that were correct, 6 answers that were incorrect, and 1 problem remained unsolved.

Table 20 provides the same information for individuals and groups working on the three problems in the second half of the investigation. Since there were 17 individuals and 3 problems,

the maximum number of solutions could have been 51. As it was, there were only 3 solutions. In 36 cases the answer given was incorrect, and in 12 cases there were no solutions. Since there were 5 groups and 3 problems, at best there could have been 15 solutions. Actually there were 4 solutions, 10 incorrectly solved problems, and 1 unsolved problem.

TABLE 19

Number of individuals	Problem 1	Problem 2	Problem 3	Result	Total
21	3	2	2	Solved	5
21	18	21	11	Incorrect	50
21	0	0	8	Unsolved	8
Number of groups					
5	3	3	2	Solved	8
5	2	2	2	Incorrect	6
5	0	0	1	Unsolved	1

TABLE 20

Number of individuals	Problem 1	Problem 2	Problem 3	Result	Total
17	3	0	0	Solved	3
17	14	6	16	Incorrect	36
17	0	11	1	Unsolved	12
Number of groups					
5	4	0	0	Solved	4
5	1	5	4	Incorrect	10
5	0	0	1	Unsolved	1

The false solutions were further analyzed. They were divided into kinds. In problem 1, the mistakes were made at different places in the series of moves taken. To make a correct solution, seven moves were necessary. No group made a mistake on the first move; 13 individuals made a mistake there. One group erred on the third and one on the fourth moves. Four individuals made a mistake on the third move and one made an error in the fifth. The results for problem 2 show about the same thing.

Thirteen moves are required to accomplish the correct solution. Individuals erred anywhere from the first to the eighth move. All groups who got as near to the final solution as the eighth move succeeded. Of the two groups solving the problem incorrectly, no errors were made until the fifth move. Problem 3 cannot be dealt with so easily on the basis of the first false move because in all but one group success resulted when once the method of transfer of the disks was hit upon.

Notes kept by the note-takers provided certain types of information pertaining to the activities of individuals within the groups. In the first place, it was found that not all in a group participated equally. Tabulations denoting which individuals made the suggestions and which ones made the rejections of ideas in the group situations were made. For example, of the four individuals in a certain group, it was found that S_1 and S_3 took the lead in the solution. S_1 made 7 suggestions and S_3 made 14. S_2 made 2 suggestions and S_4 made 3. Questions of who made the rejections and whether the suggesters of ideas made more rejections or fewer than those who made fewer suggestions were considered. There was also another question, *i.e.*, whether the suggesters rejected their own suggestions more or less quickly than those of others. In discussing the results, Shaw pointed out that an investigator might obtain interesting, meaningful, and possibly quite different kinds of results by using groups of different composition. For example, in one case the chairman might be equal in intelligence, etc., to the group. In others he might be either superior or inferior to the other members. These possibilities were offered because, in some cases, a more active member of the group than the one originally designated as the chairman took the lead in manipulating the materials during the attempted group solution. In other cases the group worked together quite smoothly, and all members seemed to contribute almost equally. In other ways, groups seemed to vary considerably in the way they worked. One group, for example, worked a solution through three times so as to be sure it had done everything right. It so happened that in the last trial an error was discovered that would have made their solution incorrect, if they had stopped at the second time over.

Summary. Shaw summarized her work somewhat as follows. She reiterated that her purpose was to compare the ability of individuals and groups of four persons in solving certain intellectual problems. Each of the problems involved a number of steps, all of which had to be right before the correct solution resulted. She admitted that the problems were somewhat removed from the life situations usually encountered.

Shaw also recognized that, although the groups were fairly well equated, they were composed of highly selected individuals as compared to the population as a whole. She concluded that groups seem more likely to arrive at correct solutions than do individuals. This was certainly the case within the framework of the investigation, but we are not so certain that this is always the case. Were problems with a number of solutions, none of which could be called absolutely correct, tried, the results might contrast with results of working problems that are mathematical (*i.e.*, have a single correct solution). Curious and inconsistent forms of compromise between individuals in the group might have resulted where solutions were normative rather than absolute. It is obvious that in such cases "matters of opinion" would have been involved and the arrival at solutions tolerable to all members of the group would certainly have involved compromises. Deadlocks might also have occurred. For example, many tasks that committees have to perform are *not* those of arriving at a mathematically correct solution, but those that will appeal to the intelligences and backgrounds of the individuals in the group making the solution or to larger groups which the committees represent. In such cases various forms of compromise and aggressiveness play roles that are certainly absent in a study such as Shaw made.

Shaw attributes the superiority of the groups in obtaining correct solutions to be caused by the rejection of incorrect suggestions and by the presence of several sources in the group for checking errors. In the investigation it was found that more incorrect suggestions were rejected by some other member of the group than by the individual who proposed them. As was earlier mentioned, Shaw concluded that not all individuals participated equally. She also pointed out that in incorrect solutions where

the steps in arriving at them could be checked, it was found that groups do not make mistakes so soon as the average individual does.

QUESTIONS

1. Discuss the major considerations involved in the general problem that is dealt with in the present chapter.
2. Describe the general experimental design of Shaw's study.
3. Describe the specific problems given to the subjects.
4. What feature did all the problems have in common?
5. What other kinds of problems might Shaw have studied?
6. What were the results in Shaw's experiments?
7. Can you think of types of problems that might have been solved better by individuals? If there are such, why might individuals surpass groups in their solution?
8. In what ways did individuals and groups differ in their problem solving?
9. Are people that make the most suggestions the most critical of them?
10. Might the results have been materially different had the groups been mixed?

CHAPTER 43

PERSONAL FRAMES OF REFERENCE AND SOCIAL JUDGMENTS

Many judgments expressed by individuals are indicative of the operation of social norms. Judgments involving social values are particularly of this kind. It has been apparent that judgments made even in psychophysical experiments in the laboratory do not escape the influence of social norms.

The acceptance of social norms of a given culture by an individual seems to come about by integration and development of systems of value and reference rather than by the wholesale acceptance of readymade fiats or precepts. Allport avers that individuals build up certain types of enduring attitudes expressing social values that pertain to classes rather than to single situations or objects. These in turn exert a directive influence in specific situations as they arise.

The matter of the *development* and *structure* of attitudes is of considerable interest in social psychology. This broad and basic interest is unfortunately relegated to subordinacy in favor of the study of highly specific attitudes and the comparison of attitudes of different groups relative to some immediate or transient issue. Such studies, if conducted properly, utilize good technique and may indirectly contribute to social science. On the other hand, they may have no more rightful place in science than any other mere poll.

Allport suggests that attitudes may be acquired in one or more of four possible ways: (1) integration of experience; (2) differentiation from general feelings; (3) traumatic or dramatic experience; and (4) acceptance of the attitudes of others.

A great many studies of attitude have been made in the attempt to understand attitude development in terms of specific factors

in the backgrounds of the persons exhibiting the attitudes in question. In the end, the conclusion has generally been that the decisive factor in attitude development is something personal and relatively difficult to get at, rather than being some so-called objective fact such as the individual's social or economic status. Status can be measured, of course, from certain tangible outward criteria, or it can be determined by the possessor's own concepts of his position. Hyman's study regarding economic radicalism and status led to the conclusion that something more than the assessment of objective conditions is necessary for an understanding in this general area. The individual's acceptance or rejection frequently cannot be predicted from knowledge of his background alone.

In the present chapter, we shall deal with Kay's study of the role that personal frames of reference play in social judgments. More specifically, this study deals with a comparison of the relation between social norms and the more personal frames of reference as applied to the evaluation of certain well-known occupations.

The specific problems chosen by Kay for study were somewhat as follows:

1. The establishment of a "social norm" by obtaining judgments regarding 12 diverse occupations with respect to five different characteristics of the men in them.

2. The determination of the influence of different amounts of personal preference in relation to the norm established by the group—the social norm.

3. The determination of the influence of different affective experiences in regard to the 12 occupations upon subjects ranking them.

4. The determination of the influence of the specific definition of the occupation on judgments regarding them.

Procedure. The main experiment had two parts. In the first part, a group of subjects was asked to rank each of 12 occupations by using five scales, each separately. The occupations were business, engineering, carpentry, law, journalism, medicine, pharmacy, music, politics, teaching, social work, and farming. The traits for which they were to be scaled were idealism,

conscientiousness, social usefulness, intelligence, and stability of character.

For the experimenter's purpose, the method of *ranking* was used, since she was interested in the relative positions of the occupations. Although these could have been deduced from a *rating method*, ranking was preferred because of the hesitancy of subjects to use the highest and lowest categories in a scale. Accordingly, the occupations would be given a relatively small spread over the scale used.

The occupations chosen included those of different degrees of interest to the subject; those that would vary in intellectual demands; and those that would vary in monetary return and prestige.

When the list was presented to the subjects, the instructions were somewhat as follows:

You are to rank several occupations on the basis of the degree to which their members possess certain traits or characteristics. *Rank one characteristic at a time.* In rating idealism, for instance, you are to put a "1" in front of the occupation whose individuals are most idealistic, a "2" in front of the occupation whose individuals are next most idealistic. Continue until you have placed the number "12" in front of the occupation whose individuals are least idealistic. *Be sure to rank each occupation for each characteristic.*

Whereas the individual scores for the various traits were found of great value for showing some of the relationships uncovered in the next part of the experiment, much of the analysis was made on the basis of a combination rating of the five reference variables (idealism, etc.). The most favorable score for an occupation, according to this scale, was 5, and the least favorable was 60. This combination score was taken to indicate an over-all summary of the subject's judgment of a given occupation.

The second part of the experiment consisted of an interview. This was held from 1 to 3 weeks after the ranking procedure. Out of the original 106 tested, 101 persons responded to the request for an interview.

The interview was designed to bring out information which would help the experimenter to understand the reference points

used by each subject in ranking the occupations. A previous, preliminary experiment had given clues as to the kind of material an interview might yield. From these clues an outline questionnaire was formulated to serve as a guide in the interview. The interview lasted about one hour, except for a few cases which ran to double that time.

The subjects were asked to indicate how many people they knew in each occupation and whether these individuals were typical members of their occupation. This part of the interview thus dealt with the subject's *acquaintance* with the occupations.

The subjects were asked what aspects of each occupation they liked and disliked most; the advantages and disadvantages of each occupation; and whether they would like to be in each of the several occupations. For this last question, one of five answers were requested ("like very much," "like," "indifferent," "dislike," "dislike very much"). They were asked whether they had been thinking of the whole occupation or some aspect of it when they made their original choices. The problem of accurate recall did not enter into this question, for the interviews had been scheduled with long enough intervals for the original thinking possibly to have been forgotten. The interview was used as a loose measure of reliability of the judgments originally made. Discrepancies between original ranking and interviews were detected and discussed. Of such discrepancies there were in the neighborhood of only a dozen out of 1,212 rankings.

A second experiment, using 60 elementary psychology students, was conducted to determine how the differently defined subgroups within the occupations were regarded. The items used as subgroups were gathered from the interviews in which it had been asked whether the individuals in these occupations were taken to be representative of the occupations as a whole or representative merely of certain subgroups within them.

The results of the experiments were dealt with in seven different ways:

1. The relative positions of the occupations were first determined.

2. Intercorrelations were made among the different scales (conscientiousness, idealism, etc.).

3. The different amounts of personal preference for the occupations were related to the judgments made by the subjects.

4. Personal experiences of the subjects were related to judgments of the occupations.

5. The combined influences of preference and experience were taken into account.

6. The subjects' understandings of (views toward) the occupations were examined.

7. Subgroups within the occupations were examined.

Results. The relative positions of the 12 occupations studied are given in Figs. 116, 117, and 118. In Fig. 116, the mean total

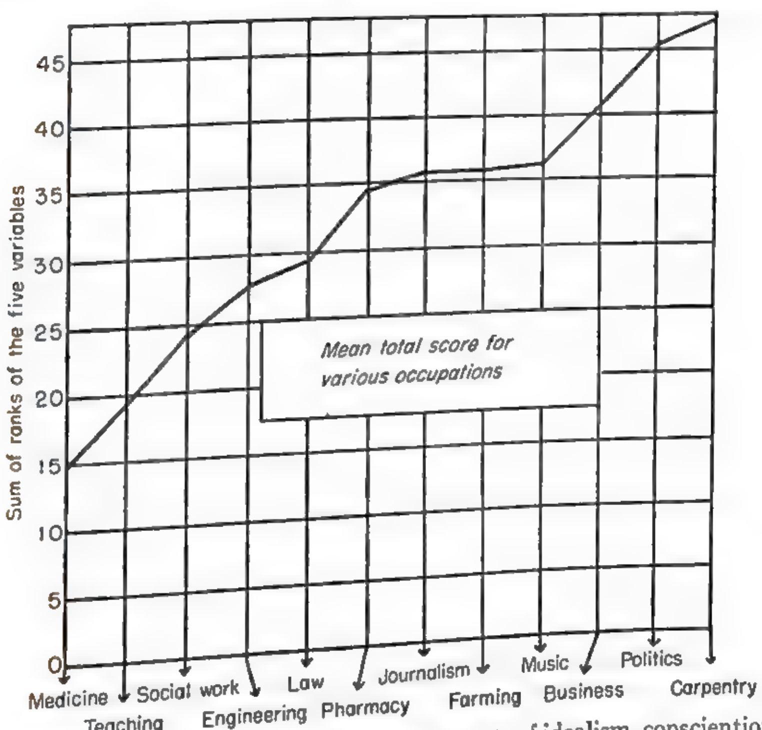


FIG. 116. Combined ratings of various judgments of idealism, conscientiousness, intelligence, social usefulness, and stability of character in Kay's study.

scores obtained by adding the ranks for each of the five characteristics dealt with are shown. Low scores are favorable and high scores are unfavorable. It is obvious then that medicine was judged most favorably and carpentry least favorably. Figure 117 indicates the ranks for the occupations with reference to three

of the five characteristics of the members in them, namely, conscientiousness, idealism, and intelligence. Figure 118 does the same for the characteristics of social usefulness and stability of character.

The rankings for the five variables were not reached in the same way by the subjects as was indicated by the interviews and

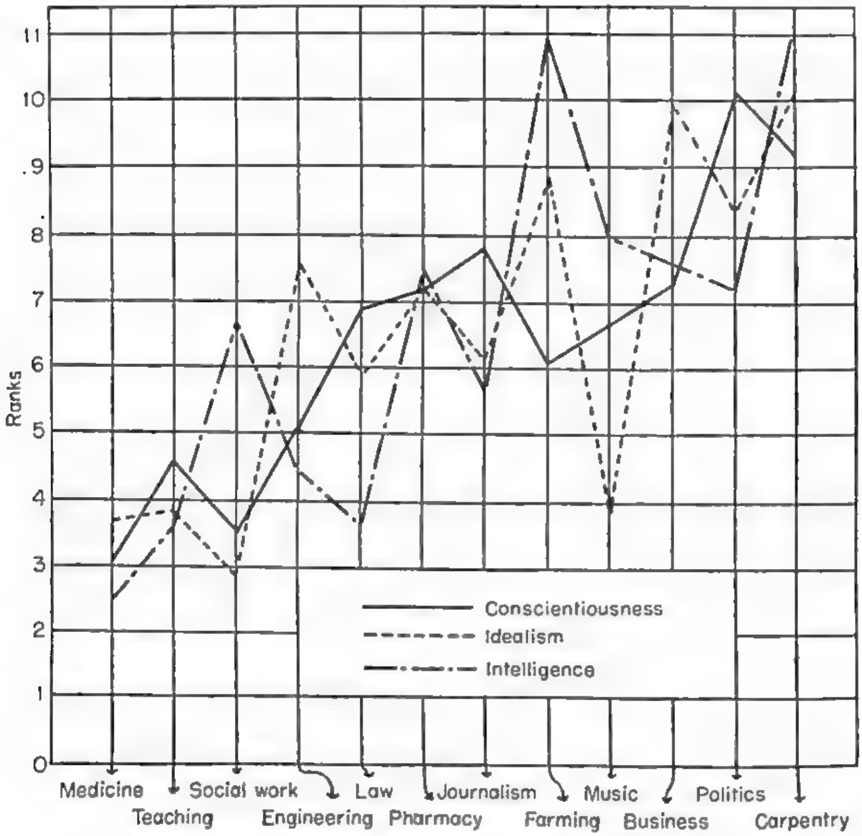


FIG. 117. Occupation ratings on three traits.

by the fact that different scales were arrived at for each of the characteristics. Farming ranked, for instance, third on the social-reference scale and eighth on the total-score scale.

The intercorrelations among the several scales varied from .50 to .95. The lowest correlation was 3.3 times its probable error. Likewise each characteristic correlated highly with the scale formed by totaling them. The range of these correlations was .79 to .86. The characteristics of social usefulness, intelligence,

and idealism, as a combination, correlated .96 with the five-characteristic scale. This result indicated that subsequent evaluation of the influence of other frames of reference will not change matters much if a scale developed from all five variables, or from some combination of them, is used.

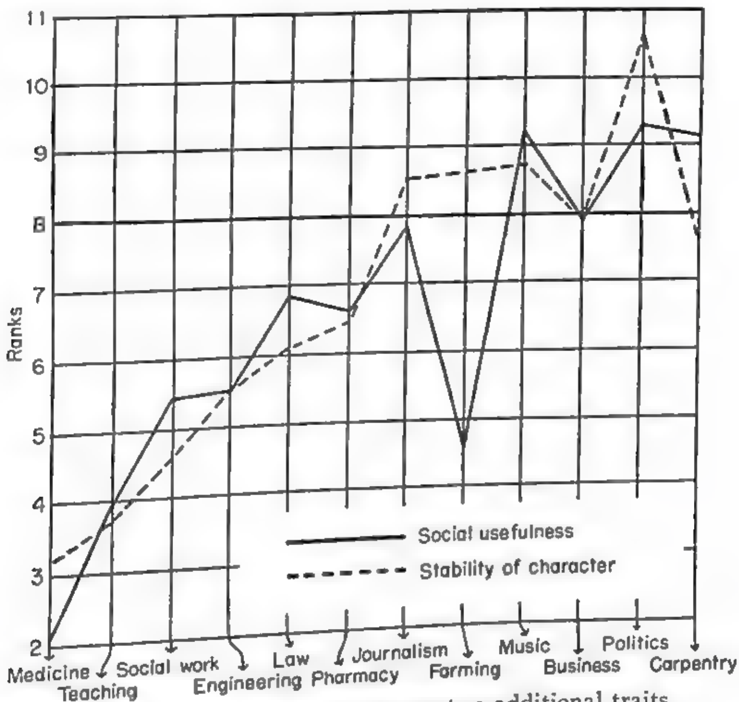


FIG. 118. Occupation ratings on two additional traits.

The next question pertained to the relation of different degrees of occupational preference and the ranks actually given the occupations. A 5-point scale was used, as mentioned earlier. There were both "likes" and "dislikes" for every occupation. For carpentry there were no "like very much" responses. It was found that liking or disliking to be a member of a given occupation was related to the subject's judgment of the characteristics of that occupation. The statistical relation was moderate even for those occupations in which a correlation was established. In 5 out of the 12, no reliable relation was found. Only in the case of music was there as much as a three-rank difference on the scales of desirability and of "dislike." But when correlations between

degrees of preference for each of the occupations and the way each was ranked with reference to idealism, etc., were made, little relation was found. Correlations between the scales for the occupations made by those who liked occupations and those who did not like them were rather high.

The relation of personal experience of the subjects with reference to the various occupations and the way they ranked the occupations was examined. For only 5 out of the 12 occupations were many experiences reported. These occupations were medicine, teaching, pharmacy, farming, and business. The other 7 occupations were put into two groups by Kay. It was very typical for subjects with positive (good) experiences regarding an occupation to rank it more favorably than the same occupation was ranked by certain other subjects (those with negative experiences). Nevertheless, subjects who had positive experience with occupations low on the rank-scale placed these occupations lower than the occupations higher on the scale were ranked by subjects with negative experiences in such occupations. Medicine and teaching, generally highly regarded occupations, were placed higher on the scale even by those having unfortunate experiences with them than were the occupations of pharmacy, farming, and business by those having favorable experiences with these three occupations. It was found that the variable, idealism, was most influenced by the affective experiences that the subjects had regarding the occupation.

The matter of the subject's perception of the occupation was studied. During the interview, the subjects were queried on what they liked most and disliked most about each occupation. The answers regarding what the subjects liked were divided into four categories: That they liked the actual job; that they liked the usefulness of the job; that they liked the people in the jobs; and that they liked nothing about the job. The "dislike" responses were, by necessity, divided as follows: "disliked nothing;" "disliked the people and/or the practices in the occupation;" and "disliked the work itself."

There was a positive rank-order correlation of .70 between the ranks of the occupation on the scales of desirability (idealism,

etc.), and the per cent of the subjects who stated that they liked nothing about the respective occupations.

The average number of individuals who designated "usefulness" as the most likable characteristic of the first six occupations was more than twice the average number who indicated usefulness for the last six occupations.

When the subgroups within the occupations were rated by a second group of 60 subjects, there was a correlation of .92 between their results and the earlier ratings of the occupations as wholes. Quite similar ratings were found between subgroups having quite closely related functions, as for example, between corporation lawyers and corporation directors.

Kay concluded that, in the experiments, she was dealing with three frames of reference. One of these was the frame which affects the whole group and is called a social norm. The second was the socioeconomic position of the individual subjects. The third factor was the highly personal and individual frame of reference of each subject.

QUESTIONS

1. According to Allport, in what ways may attitudes be acquired?
2. Is the consideration of the socioeconomic conditions under which an individual lives a good basis for predicting attitudes?
3. What did Hyman's study show in this respect?
4. What was the objective of Kay's study?
5. What were the specific problems chosen?
6. What kind of a scale of measurement did Kay use in her investigation?
7. Why was this scale preferred to some other one?
8. To what use was personal interview put in Kay's study?
9. Compare the subjects' performances in ranking the five variables used.
10. Summarize Kay's findings.

PART X

LEGAL PSYCHOLOGY

CHAPTER 44

FORMATION OF OPINION BASED ON LEGAL EVIDENCE

Despite the extreme importance to society and to the individual the results might have, little or nothing has been done experimentally to understand how the individual functions in a court trial. A number of persons are faced with forming opinions in a trial. The task of some of these persons is to see to it, if possible, that the net opinion is in favor of the defendant. The task of others is just the opposite. Still others must follow the presentation of evidence in order to form opinions of their own. The problem of ultimate formation of opinion is involved first in the nature of the indictment, then in the selection of a jury, the choice of evidence to be revealed, and the evidence to be omitted, and, finally, in the way the evidence is verbally interpreted and summed up by the attorneys.

If you keep in mind that it is as much the defending attorney's rightful task to bring the case to an acquittal as it is the prosecuting attorney's job to gain a conviction, you will not be surprised at the complexity of a court case. You will readily see that it is very likely that most laymen are in no position to sift the evidence and render a rational opinion. Yet it is steadfast opinions they must form and do form. Added to the complexity of the courtroom portion of the trial are the deliberations of the jury where the matters already possibly not too sure in the minds of some of the jurors are further garbled.

Since many trials are life-and-death matters for some of the individuals in them, considerable study ought to be spent on understanding what is involved in forming opinions. It ought to be recognized that a jury trial is a kind of test of the partic-

ipating jurors as well as a trial of the defendant, even though it is not conventionally thought of in this way.

In the literature, several jurists have made *logical* analyses of proof. But from what is known of the psychological functions of comprehension and thinking, Weld and Roff assert that it is very likely that the operations involved in the formation of opinion are not so simple as a logician's analysis would lead one to suppose.

Weld's and Roff's Experiment. In the present chapter, we shall deal with a recent study that Weld and Roff made in this area. The experimental conditions which they set up are not intended to be fully identical to those in a court trial. The essential plan which they followed was that of presenting evidence, bit by bit, first by one side of the case and then the other. They stopped the flow of evidence at various stages and obtained opinions or judgments at these points. The authors also eliminated the customary summing up of the evidence as by the trial lawyers and also the discussion of the evidence in the jury room. Since the evidence used in the experiment was only from a report of a case, the actual presence and behavior of the witnesses were also eliminated. Nevertheless, it was felt that the experimental procedures used would throw some light upon the general process of opinion formation, and to that extent provide some understanding of the manner in which a jury functions.

Method. For the experimental material, a famous bigamy case, the Thomas Hoag case, was selected. The testimony as found in the report could be read in an hour. In the order found in the report, the material was such that sure opinions could hardly be formed before the conclusion of the report. In the experiment, the testimony was read slowly and clearly to groups of college students (upper classmen) who were preparing to be lawyers. The same group was used only once, and the experiment lasted several years. Prior to the reading of the report, the data such as the name and alias of the defendant, the names of the alleged wives, the essential dates, and the names of the places involved were all listed on a blackboard. The testimony was divided into sections or installments designated by Roman numerals. At the conclusion of the reading of each installment, the reader stopped and required all the listeners to record their

judgments as to guilt or innocence of the defendant on a 9-point scale. The scale was substantially as follows:

1. Certainty of innocence.
2. Strong belief in innocence.
3. Moderate belief in innocence.
4. Slight belief in innocence.
5. Doubt as to whether innocent or guilty.
6. Slight belief in guilt.
7. Moderate belief in guilt.
8. Strong belief in guilt.
9. Certainty of guilt.

For the first installment of the trial, the indictment was read. Naturally, this was not testimony, but it was proper to see whether or not the mere indictment might have an effect upon the opinion of the subjects.

The final installment, the thirteenth, was the reading of the verdict of the actual trial. This installment, too, was not evidence, but was used to determine whether an effect, if any, would be made upon the final judgments of the subjects.

When all the installments and their consequent opinions were completed, each subject was required to write a short statement regarding the evidence that was most influential in forming the judgments. From year to year the evidence was presented in different orders. This was the only variation involved.

The whole case report is too long to quote completely, so we shall give the gist of one or two installments.

Installment I—Indictment. The prisoner, supposedly Thomas Hoag, formerly of the County of Rockland, laborer, alias Joseph Parker, now of the City of New York, teamster, did on the 8th of May, 1797, at New York, lawfully wed Susan Faesch. The said Thomas Hoag, alias Joseph Parker, afterward, on the 25th of December, 1800, at the county of Rockland, his lawful wife being alive, did feloniously marry and take to wife, one Catherine Secor. To this charge, the prisoner pleaded not guilty.

Installment III—Prosecution. John Knapp testified that he knew the prisoner for 2 years beginning in 1800; that the prisoner was then in Rockland County and was known as Thomas Hoag;

that he saw him daily for 5 months and was at his wedding; that the prisoner had a scar on the sole of his foot as a consequence of stepping on a drawing knife; that Hoag showed him the scar; that he, the witness, was convinced that the prisoner was Thomas Hoag.

Catherine Conklin (formerly Catherine Secor) testified that she came to know the prisoner early in September, in 1800, after he came to Rockland; that he then went by the name of Thomas Hoag; that she saw the prisoner constantly; that soon after their first meeting, the prisoner began to keep company with her; that finally on December 25 he married her; that they lived as man and wife till late in March, 1801; that then he left her; that she did not see him for 2 years; that on the day of leaving her he seemed to want to tell her something that was on his mind, but was persuaded not to by a companion who professed to be his brother; that the prisoner, to the very last, was a kind, affectionate, and attentive husband; that she was as well convinced as she could be of anything that the prisoner at the bar was the same person who married her; that she then considered him, and still does consider him, the handsomest man she ever looked upon.

As was stated earlier, the 13 installments were not given in the same order to all groups of subjects. Figure 119 indicates the way the opinions formed and changed as the evidence was given in the first order (I to XIII consecutively). In the graph in this figure, as well as in the three subsequent figures, the ordinate represents the degree of certainty of judged innocence or guilt, starting with judgment 1 which indicates "certainty of innocence." The abscissas in each of the graphs indicate the installments and their order of presentation to the subjects.

The graphs present not only the values of the mean judgments for the group but also the quartile deviations. Upon the reading of the indictment, the subjects were far from neutral. One subject was convinced of the prisoner's guilt, 4 of the subjects held a strong belief in that direction, 6 held a moderate belief in the same direction, 22 held a slight belief that the prisoner was guilty, 15 were in doubt (the position that all ideally should have held), 1 had a slight belief in Hoag's (or Parker's) innocence, and 1 had a moderately strong belief in the defendant's innocence.

Since the median at the time of the indictment was 5.9, with a quartile deviation of 0.6, the group was influenced somehow to believe slightly in the prisoner's guilt.

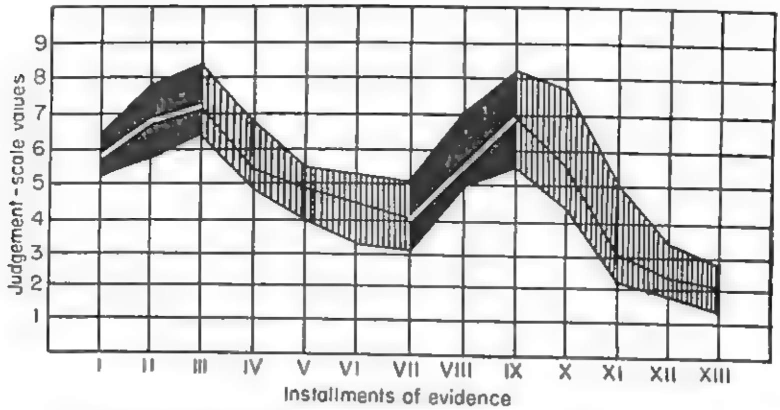


FIG. 119. Relation between judgments and installments of evidence in experiment 1. The central line indicates the median of all judgments. The width of the band indicates the second and third quartile deviations. Black portion indicates effect of evidence for the prosecution; the barred areas, the effect of evidence for defense. (*Weld and Roff. By permission of Amer. J. Psychol.*, 1938, **51**, 618, *Fig.* 1.)

The order of installment presentations shown in Fig. 120 was changed to I, II, III, IV, V, VI, VII, X, XI, XII, VIII, IX, XIII. Here again, the group starts out predisposed to believe in the prisoner's guilt. The mean is higher than for the first order of

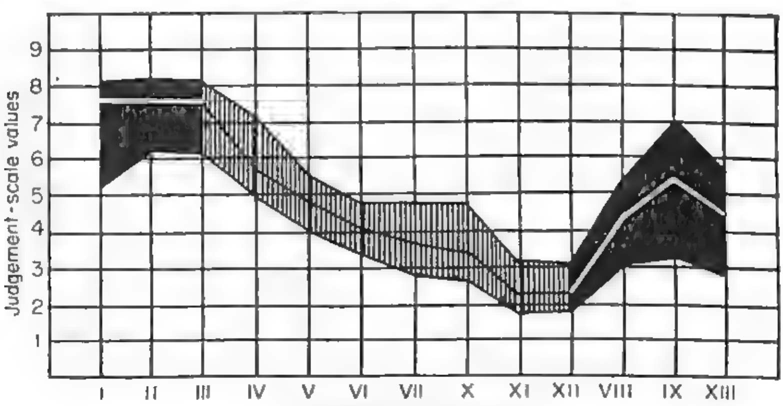


FIG. 120. Relation between judgments and installments of evidence in experiment 2. (*Weld and Roff. By permission of Amer. J. Psychol.*, 1938, **51**, 620, *Fig.* 2.)

presentation, but the first quartile deviation is about the same. In the present order, the trial opens with the indictment and prosecution and then soon begins with the defense. The defense testimony is all given, except one installment at the very last that was really no testimony. It was the verdict. As a consequence the presumption of guilt gradually diminishes and turns into belief in innocence. When the prosecutor takes the stand, however, the median rises again toward belief in guilt and finally just barely passes the doubtful point (point of neutrality, at judgment 5.).

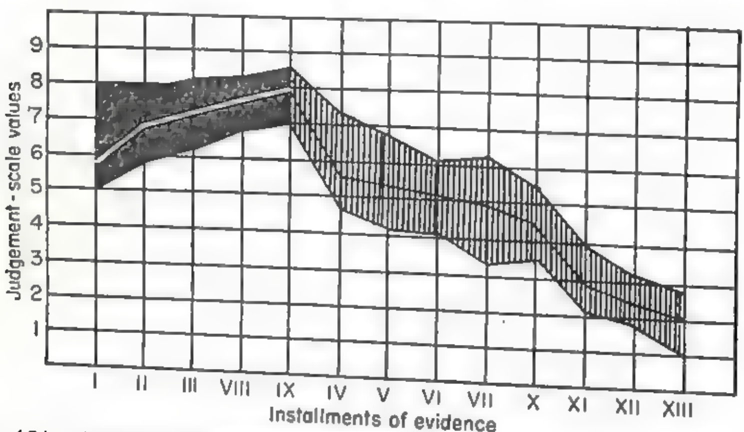


FIG. 121. Relation between judgments and installments of evidence in experiment 3. (*Weld and Roff. By permission of Amer. J. Psychol.*, 1938, **51**, 620, *Fig.* 3.)

The installments as shown in Fig. 121 are in the usual legal order (aside from rebuttals). All the evidence of the prosecution is presented first (installments I, II, III, VIII, IX), and all of the defense is then presented (installments IV, V, VI, VII, X, XI, XII, XIII). At the end of the case for the prosecution, the median was about 7.9, representing a strong belief in the prisoner's guilt. At the end of the defense testimony the median judgment dropped to 2.0. This final result was the same as in experiment 1, shown in Fig. 119.

In experiment 4, shown in Fig. 122, the order of installments was again changed to become I, IV, V, VI, VII, X, XII, II, III, VIII, IX, XI, XIII. Again the pattern of sway from belief in innocence to belief in guilt was altered. In the present order, the

testimony for the defense followed the indictment and continued until it was all presented except installment XI, the showing of the prisoner's foot to the jury. The prosecution was all presented, and then the foot was shown. To help to make the matter clearer to you, it should be pointed out that when the foot was shown, no scars whatsoever were found, as the jurors might have been led to expect by the prosecution. You should also know that the last installment, the verdict, was that of "not guilty."

It is noteworthy, then, that despite the verdict of "not guilty" and the lack of scars on the foot, the order of presentation used in

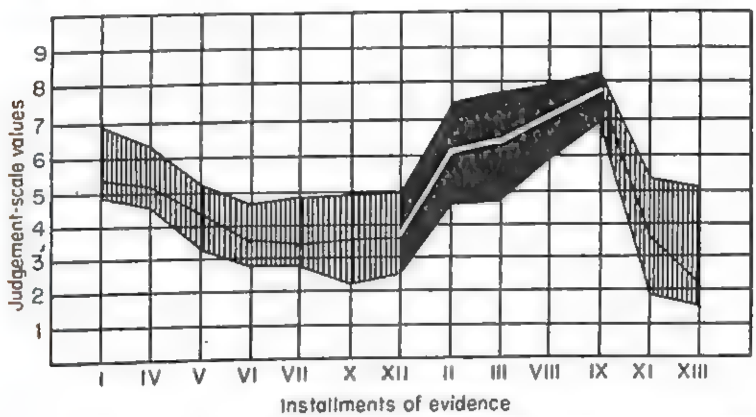


FIG. 122. Relation between judgments and installments of evidence in experiment 4. (*Weld and Roff. By permission of Amer. J. Psychol.*, 1938, 51, 621, *Fig.* 4.)

experiment 2 resulted in a final median score of 4.4 just a little less than slight belief in innocence or almost neutrality.

Weld and Roff also analyzed the serial judgments of 12 of the group of subjects who participated in experiment 4. These were taken as representative of the differences among subjects in all the experiments. The subjects were labeled by letters of the alphabet.

A, J, and L held a position of doubt until the beginning of the prosecution's evidence. As was afterward admitted, A and J did not follow the instructions of the experiment. Regardless of the evidence, they thought they should maintain their doubt until all the evidence had been presented. Despite this, neither one of them was able to do so. Curiously enough, J, at the beginning

of the prosecution's evidence, concluded that the case was one of mistaken identity, and all evidence from then on only varied his certainty regarding the prisoner's innocence. L was sincerely doubtful with a slight belief in favor of guilt. This attitude was enhanced by the evidence for the prosecution until he finally became sure of the prisoner's guilt. The absence of the scar on the prisoner's foot led two of the subjects to conclude innocence, but for L, it diminished but did not erase the judgment of guilt.

Seven out of the 12 were more or less gradually influenced toward belief in innocence as the testimony for the defense accumulated. All of these subjects moved toward guilt as the testimony for the prosecution began, and all but one of the seven believed to some degree in the prisoner's guilt until the foot was shown as evidence.

One subject out of the 12 was so convinced of the prisoner's guilt by the mere indictment that he never got to less than judgment 6 (slight belief in guilt). In his way of thinking, the evidence for the defense was self-contradictory, and this kept the way open for the possibility of guilt. This possibility was strengthened by the evidence presented by the prosecution.

When the judgments for all 12 were compared at installment IX, all but 2 individuals were moderately strongly convinced that the prisoner was guilty. When the foot was shown in the next installment (XI), 4 subjects changed their judgment to innocence, 3 reduced their judgments to doubt (5), 2 only became less certain of guilt, and 1 shifted to judgment 4, a slight belief in innocence.

It was obvious that B, E, and H were easily influenced by previous judgments, whereas D, F, G, and I were more cautious in their opinions and shifted only a step at a time, if at all, and ended only in the middle of the scale—at doubt.

As you will recall, part of the experimental procedure was to determine from each subject what fact was most significant in influencing decision. As it turned out, scarcely a piece of testimony in the case was not judged by someone or another to be crucial. For most persons, however, the absence of a scar on the prisoner's foot was the most telling argument. Some thought that Catherine Secor, who had been Hoag's wife, could not have been mistaken in her identity of the prisoner.

The conclusions that the authors came to with regard to the formation of opinion ran something as follows: Under the conditions of the experiment, the subjects are faced with a series of evidential facts. The subjects, in effect, evaluated them as attitudes, the probability of the occurrence of the facts, and whether the allegations presented were reasonable or not. Finally, the prisoner's plea was taken into account. From all this, the determination of opinion appears to be far more immediate than you might suppose. There seems to be no evidence of a reasoning process involved or that opinion is syllogistic. It is an immediately felt affair.

The technique of Weld and Roff is certainly suggestive of a great deal more that could be done in the way of penetrating the process of judgment and the formation of opinion.

QUESTIONS

1. How does the present study in legal psychology differ from the more common type or types?

2. Point out the essentials of experimental design in Weld's and Roff's study.

3. Where on the 9-point scale should the judges or jurors *logically* start out?

4. Actually, where did they begin?

5. What might be the confusing features of the experiment to the individuals used as jurors?

6. Suggest modifications of procedure that could be used in this study.

7. What variable or variables were used by Weld and Roff?

8. Compare the results accruing from manipulation of the order of installments of evidence.

9. What was discovered regarding the facts that most influenced decisions?

10. Is opinion based upon reasoning? Justify your answer.

PART XI

CHILD PSYCHOLOGY

CHAPTER 45

A STUDY OF YOUNG CHILDREN'S PLAY FANTASIES

One of the commonly made divisions in psychology is that of child psychology. There are times when it is spoken of as genetic psychology, but more properly, genetic psychology is a broader form of developmental psychology including the study of the various stages in the life of subhuman as well as human species. Child psychology has added significance beyond what it would have as a part of developmental psychology in general. This stems from the fact that knowledge regarding the growing child is the basis upon which educational and other systems of dealing with children are, or should be, designed. Naturally, attention to the developing individual provides us with kinds of knowledge about adults that would be difficult or impossible to obtain more directly.

Our main concern is in seeing that the study of children is conducted on a thoroughly sound experimental basis. This involves determining first what it is about children's behavior that should be studied. It is obvious that the experimenter cannot take the child into an artificial situation and apply instructional controls, as he often does in the case of adults. Another of the chief features in the study of children is the avoidance of depending upon the child's having to answer direct questions, and thus having to become his own evaluator through verbalization. Whether in dealing with children or adults, this dependence upon answering questions is too near to the analogous case of a patient having to make his own medical diagnosis.

The object, then, in studying children experimentally is to lay conditions under which the children reveal themselves by what they do and the incidental remarks they make. In one sense,

this is quite a direct method, and in another, it can be said to be one of indirection because the actions must be subjected to interpretation. The acts are only representative or symbolic of the principles in terms of which ultimate descriptions and understandings are to be couched.

Although we may apparently be inconvenienced because we cannot get at the underlying mechanisms of action more directly, the possible modes of investigating children may hold certain advantages. Children may quite fully and easily be led into situations in which many features of their behavior become transparent. Children particularly demonstrate this possibility in cases in which they are led to play roles. Involved in role playing is what may be called fantasy, or the child's exercise of imagination regarding type-situations. The situations concerning which the child's imagination is involved are those about which it has had experience and to which some significance is attached.

In this chapter we are going to deal with Bach's study of children's play fantasies. Ordinarily one of the most difficult aspects of a study of this sort is the great amount and variety of fantasy material that is manifested and may need to be recorded. Much of this was overcome, for instance, in the procedures used by Bach. Bach reduced the notations to manageable proportions by standardized procedures. He pointed out, in the report of his studies, that the success of an approach to the study of personality does not depend upon the extent to which it yields a complete picture of the total individual personality, but rather upon its usefulness in the solution of diagnostic and predictive problems. Bach's approach was based upon two considerations: (1) the waiving of some of the spontaneity of the children's fantasy expressions in order to permit experimental standardization, the object of this standardization being the reduction of variety in fantasy response; and (2) the forfeiture of obtaining comprehensive data on each child's behavior during the experimental play session so that information in a few carefully selected, fantasy-activity dimensions might be obtained. Bach's work was as much an attempt to establish some of the experimental conditions necessary for studying the type of behavior in question

—fantasy such as is involved in children's doll play—as it was to gain specific information. He wanted his procedure to be one by which everyday social behavior could be predicted from clinical or experimental measures. Accordingly, it was necessary, as part of his procedure, to ascertain some of the everyday behavior variables related to different types of fantasy activity.

In order to compare the behavior of the children in the experimental play sessions with their outside behavior, six experimental rating scores were constructed, each for a certain area of the children's everyday activity. These areas pertained to (1) eating, dressing, keeping clean, etc.; (2) the tendency to accept guidance during play activities; (3) the children's emotional attachment to their teacher; (4) the children's social success in contact with other children; (5) the children's activity preferences, such as block play, water play in basin, etc. This test was divided into indoor and outdoor activities. The last rating scale in this group (6) concerned the children's imaginative play.

In Bach's plan, the following characteristics were demanded:

1. Equal opportunity for all subjects to express themselves with regard to similar matters had to be given. This was necessary so that each subject could be assigned to a position in each of the dimensions to be dealt with in the experimental procedures.

2. The degree of prediction from fantasy responses to overt social behavior of the child should be capable of investigation so as ultimately to be known.

3. Fantasies (accessible to measurement) that could reasonably be supposed to have their antecedents in social adjustment outside the experimental situation should be elicited. That is, there should be two situations with relationship to each other, each capable of being measured. One situation would be the experimental ("clinical") one; the other, the outside (everyday) situation. The experimental situation should be devised from what is known about the outside social situations.

Subjects and Conditions. A stylized dollhouse, copying the preschool building and its contents (furniture, etc.), with which the children were familiar, as constructed as play equipment for the experimental periods. The subjects were youngsters enrolled

at the Iowa Child Welfare Research Station. Fantasies on one general school theme owing to the common background in school life were used. It was supposed that such children would be less variant in their range of experiences at school than at home. This was expected to render less difficult the interpretation of individual differences in fantasies in regard to school than to home or elsewhere. This, of course, did not rule out the occurrences of fantasies having to do with home life, for there were motivational factors that were certainly common to both situations.

The subjects were 55 in number, and ranged from 34 to 64 months in age. Fifteen of these subjects were used for an extensive initial procedural investigation lasting 9 months. Of the other 40 children, 5 did not remain in school to fully complete the experimental schedule. Twenty of the remaining subjects were of the younger preschool group (group III), and 15 were from the next age group (group IV). Group III met in the morning and group IV in the afternoon. Although they met in the same building, they had different teachers. Four of the teachers participated in the study.

For each subject, the age, sex, mental age, IQ, Smith-Williams vocabulary-test scores, and parents' occupation were noted. The following list indicates the number of children who had fathers with the occupations indicated.

6	Physician	4	Sport coach
17	Academician	2	Career officer
6	Businessman		

The IQ's of the children ranged from 89 to 147, only one of the youngsters having an IQ below 102. The average was 123.4 with a standard deviation of 12 points. The mental ages ranged from 39 to 81 months, with an average at 62.8, and a standard deviation of 10.3 months.

Experimental Procedure. Each child was given four similar play sessions of about 20 min. each. Generally 3 days intervened between sessions. Fifteen subjects were taken before their rest or nap, and 20 subjects following it. The experimental session was entered into through a preliminary conversation. The experimenter approached the child during his free play activities

and led up to the question whether it would like to have a turn playing with the dollhouse. The experimenter noted the activity in which the youngster was busy at the time and rated its reaction to the invitation on a 6-point scale—itemized as follows: eager acceptance; willing to come; matter of fact; neutral; hesitant; refusal to come. After receiving acceptance, the experimenter took the child across the street to the building containing the dollhouse. This was not an artificial procedure, for the children were used to being taken away from routine school activities to "play games."

The child was taken into a large, light room, devoid of furniture and supplied with a one-way viewing screen. This room contained the dollhouse previously mentioned. Upon entering the room, the experimenter pointed out the house, which was covered with a cloth, and said that the child could pull the cover off and see what was under it.

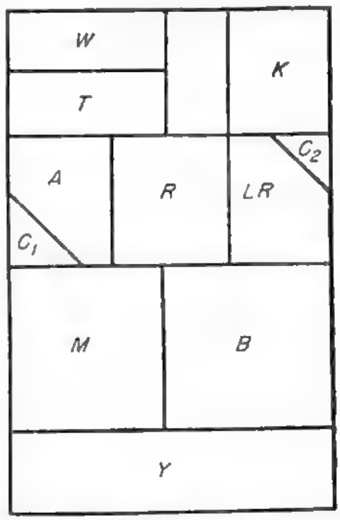


FIG. 123. Plan of dollhouse used by Bach in studying play fantasies of young children. (*Bach. By permission of Psychol. Monogr. and Amer. Psychol. Assn.*)

It was typical for the child to comply readily. It was then allowed a few minutes to look the dollhouse over. The child was invited to sit down on one of the throw rugs, and the experimenter sat down with his back to the viewing screen so that it was possible for the child to sit so as to be visible from the viewing screen. The dollhouse was quite like the rooms in the child's real preschool (see Fig. 123). It had a music room *M*, a playroom *B* with blocks in it, an art room *A*, a dining room *LR*, a washroom and toilet *WT*, a kitchen *K*, and two large closets (C_1 and C_2). The playhouse was without a roof so that the dolls could be arranged or made to perform to suit the child. The house (20 by 30 in.) and the dolls were small so that almost everything could be easily encompassed within the visual field of the child at all times.

The initial introduction to the house was without the dolls in it. This phase of the introduction was completed when the function of every room and its capabilities for doll play had been pointed out to the child, and after all the "reality-unreality" perplexities had been explained to it. The function of this extensive introduction was (1) to accomplish early satiation of nonthematic interests, so that later no interests of this sort (interests in furniture, etc.) would disrupt thematic activity; (2) to provide all subjects with the same spatial understanding, so that failure to elaborate on some subtheme could not be credited to the child's not knowing certain features of the doll's environment; and (3) to facilitate imagination ("making up something").

After the doll-less introduction the experimenter paused to give the child a chance to ask for the dolls that would be the "teacher" and the "children" for the "play preschool." Then the experimenter announced the teacher's arrival. Following this, the child frequently spoke up and asked where the children were. As an answer to this question the child sometimes made some suggestions of its own. Possibly a hunt through the closets of the room would ensue to find the "pupils" (dolls).

The three dolls that were to be the pupils were introduced one by one by the experimenter as he withdrew them from his pockets in keeping with the conversation between himself and the child. Thus the dolls tended to take on personalities of real children. After the three dolls had been introduced, some children asked whether there were any more pupils coming, but they were assured this was a small school and three would be all.

Usually, the child's change of role from happy, interested onlooker to active creator of dramatizations was quite smooth and easy. The task was finally set by the experimenter when he told the child to go ahead now and play with the dolls in any way it wanted to. The child was told to do anything it would like to.

The experimenter did not assume an entirely passive role during the play session after he had once told the child to go ahead and play as it pleased. Various kinds of so-called verbal stimulation were necessary. These fell into the categories of (1) support for fantasy, (2) stimulation of the child to "identify," (3) reduction of anxiety, and (4) facilitation of observation. In category 1 the

experimenter might demonstrate a short stereotyped theme showing how to make the dolls perform. In category 2, suggestions regarding whom the dolls looked like, which was the prettiest, etc., were offered. In category 3 anxiety reduction, personal friendliness, was displayed by the experimenter. Verbal sanctions were given when the dolls were made to perform unorthodox acts. The content of certain acts of the child once begun, but suddenly self-checked, were anticipated and endorsed. Certain "unorthodox" acts, such as having one doll hit another, were demonstrated in certain cases. To facilitate observation, the child was urged to have the dolls do something rather than for the child simply to say something. In the case of unclear (ambiguous) doll acts, the experimenter asked what was meant and had the child speak loud enough so that the person (observer) behind the screen could hear.

It is obvious, of course, that since each child was not treated exactly alike, that the generalizations which Bach could legitimately make were affected. It is not certain how well this factor was taken care of. Our interest, however, is to acquaint ourselves with the general method and attempt to think of ways of standardizing the situation or else of evaluating the experimenter's performance in relation to the child's.

Data. Three principal classes of events were noted: (1) Thematic activity, (2) nonthematic activity, and (3) experimenter's injected activity as stimulation. The term "thematic" described the activity in which the dolls were made to interact with each other by the child—play in which the dolls took on roles and acted accordingly. It involved largely what the dolls were made to do to each other by the child.

Thematic activity included both verbal description of activity of any of the dolls and actual manipulation of them in roles within the dramatic situation.

Bach stated that the rooms in the playhouse, the yard, and the surrounds gave 11 story backgrounds with which doll activity or their experiences could happen.

Reliability. The reliability of the observations was expressed in terms of per cent agreement between two independent observers who recorded what they saw. The reliability in the

various categories ranged from about + 71 to about + 96 per cent. The values were based upon comparisons involving 64 2-min. records, simultaneous but independent, by the two observers on six subjects made before observations on the experimental group were begun.

For example, the agreements involved such particulars as in the following case. Both observers in one instance agreed that within a given 2-min. period the same doll was involved, that the doll was the receiver rather than the originator, that the doll's act was one kind of nonstereotyped performance, and that this act was not a sustained but a brief thematic affair belonging to a given subtheme. A number of findings were obtained by the technique briefly described. Among them were the following:

1. The children (assumed to be normal) evidenced extremely aggressive fantasies. Although the differences in the fantasies were great, more than three-quarters of the fantasies copied realistic conditions in the individual child's everyday surrounds.

2. There was an extreme contrast between the kind and amount of fantasy in boys and in girls. The girls were more productive, possessed a greater nicety, and more stereotypy. The boys evidenced more thematic aggression.

3. The group of children subjected to the more extended routine (greater rest routine, and thus supposedly developed frustration) elaborated the *rest* theme significantly more frequently and with more aggressive expressions of it than did the other group having only a short rest period.

4. Data regarding the children's known actual social behavior and adjustments were compared with the fantasy measures obtained.

The following conclusions were made by Bach:

1. Compliant children manifested more elaborate fantasies regarding school. Resistive children manifested stronger positive transference relationships to the experimenter.

2. The rate of gradual decrease (in four play sessions) in stereotyped fantasy was a means of distinguishing well-adjusted children from poorly adjusted ones.

3. Both the outwardly aggressive and destructive child and the one who in daily behavior is neither, manifested the same great

amount of aggression in fantasy. The *normally* aggressive child exhibited less thematic aggression than the two extreme types of children just suggested.

4. Children exhibiting quite stereotyped fantasies stayed less involved emotionally than did those children who exhibited a greater relative amount of nonstereotyped thematic activity. Such children were also of the more compliant type in everyday life.

5. Children who showed thematic aggressions against the teacher were in everyday behavior decidedly less compliant to teachers than were those who showed no such fantasy aggression.

6. The children developing any noticeable elaborations of sexual and toilet themes were definitely younger in mental age than those not making such elaborations. The other children brought out a greater amount of chasing and hiding in their themes.

7. Children greatly involved emotionally in their fantasies in the experimental play periods were more affectionately related to teachers than those less emotionally involved.

8. There were differences in amount of "identification" during the play sessions, and those who identified their dolls as real people including themselves manifested a greater correspondence between fantasy behavior and everyday behavior.

We do not have opportunity here to develop an extended discussion of the methods, aims, means-to-the-end, and interpretations developed by Bach. To say the least, he demonstrated a general methodology that appears to have certain possibilities for getting at childhood personality structure, etc., that has not heretofore been too well penetrated. The scheme tends to put young children into an active situation in which one of childhood's common modes of expression (make believe) is utilized in a standardized, insightful way. To say the least, the technique should serve as enlightening and as a stimulus for you to think for yourself in this direction.

QUESTIONS

1. What two procedural considerations is Bach's study of children play based upon?

2. What data were obtained regarding the children with which to

make comparisons between everyday behavior and behavior in the experimental sessions?

3. What were the three chief specifications of Bach's experimental design?

4. What equipment was necessary for Bach's study?

5. Describe the actual procedure used in the study.

6. What role or roles did the experimenter play in the procedure?

7. Were all of Bach's subjects treated alike in the experimental procedure? If not, to what use can we put his study?

8. What is "thematic" activity?

9. What differences were detected between boys and girls?

10. What differences were brought out between the more aggressive and the less aggressive children?

PART XII

CLINICAL PSYCHOLOGY

CHAPTER 46

DELINQUENCY AND AFFECTIONAL RELATIONS BETWEEN PARENTS AND CHILDREN

Every phase of psychology is potentially a part of experimental psychology. Whether a given phase or field actually is now, or ever will be, experimental depends solely upon the ways in which problems in the area are conceived and studied. There is a way of doing genuine experimental work in each field. Whether it is found depends upon the men in the field.

It is the purpose of this section to present some examples of experimental clinical psychology. Much of what is nowadays called clinical psychology was formerly better known as abnormal psychology. Actually, clinical psychology is the psychology that is applied in clinics or hospitals where the mentally distressed or ill are treated. Since the "mentally ill" and the "abnormal" are synonymous, clinical psychology is the psychology that pertains to the abnormal. It is abnormal psychology in practice. If a difference is to be maintained between abnormal and clinical psychology, the latter is *applied* abnormal psychology.

On the other hand, if clinical psychology is to be considered as merely applied psychology, it can scarcely aspire to being *experimental* in the full scientific sense of the word. Therefore what we are dealing with here as clinical psychology must be the experimental psychology of abnormal behavior.

Our first problem in clinical psychology will pertain to juvenile delinquency. The basis (or bases) for juvenile delinquency has been a matter of discussion and diversity of belief. Of the various opinions, two main lines of thought have been expressed. One of these attributes the causes of delinquency to society, and it is therefore to be called a sociological concept. The other attributes the origins of the trouble to factors that lie within the individual

and accordingly may be called a biological concept. Biological concepts differ in terms of what features of the organism are involved.

Of the sociological studies on delinquency, we have that of Shaw and McKay which concluded that in several large American cities the incidence of delinquency is highest in areas in which there is architectural decay, declining population, and disintegration of neighborhood organization. Poor housing and congestion were said to be only expressions of these characteristics. Juvenile delinquency is regarded as the manifestation of social disorganization. The authors would suggest that physical rehabilitation of slum areas, improvement of social and economic conditions, and the formation of a community containing certain cohesive properties would solve the problem.

Accordingly, the treatment of the problem lies in doing something to groups that can take the form of physical betterment. Perhaps the argument lies not so much in distinctions between helping groups as groups, and many people as individuals, as it does in seeking out the ways in which groups as groups can be dealt with as in contrast to individuals as individuals. Group help generally takes on the form of material aid, whereas help for individuals is seldom, if ever, considered seriously in material terms. Therapists generally are dealing with personality mechanisms, etc., that cannot be corrected with dollars and cents. This distinction in the means whereby groups and whereby individuals are helped is possibly an artificial one and fairly unfortunate, but nevertheless it is a conventional one. Furthermore, it is a distinction that is involved all too often politically.

It is to be admitted that there is a portion of the delinquent population that can be comprehended in terms of sociological theories. As for example, there are children who accept the standards of a home and neighborhood that, according to some definitions, are delinquent, and as a result the children become delinquent. Sociological theories do not get at the reason why many children in a delinquent area do not become delinquent. They do not explain why one sibling becomes delinquent and another one does not. Such theories do not account for delinquency in communities and homes enjoying comfortable social

and economic status. Some studies, in fact, have found that goodly numbers of delinquents come from "good" homes. Grimberg found that most of the delinquent girls studied came from respectable families and comfortable homes. The material amount of involvement of "good" environment occurs despite the fact that not all delinquency arising in such circumstances comes to the attention of civil or social authorities. Incidence of delinquency in better residential areas is greater than any statistics would indicate. Thus the community, in terms of its conventional physical and social assets, is not to be considered the single crucial factor.

Other authorities put forth different ideas regarding the basis of delinquency. With the advent of the mental-testing era, men began to believe in inferior intelligence as the cause of delinquency. Goddard was very influential in this direction and pointed to low-grade mentality as the greatest single cause of delinquency. It came to be recognized that the correlations between mentality and delinquency were not great and were distorted in favor of the mentality theory, owing to the probability that the higher grade individuals escaped detection more often, etc.

In other ways, heredity was also invoked to account for delinquency. Such ideas received certain setbacks. During the prohibition era, we turned out to be virtually a nation of delinquents. In this example, the heredity idea of delinquency was not invoked.

Our specific considerations on the problem of delinquency will be in the form of describing the study of Zucker, which dealt with the relation between delinquency and the identification of the child with its parents.

The hypothesis which Zucker formed was that the affectional identification of most delinquent children with their parents is either inadequate or absent and is a condition functionally related to the degree to which such children do not take to themselves the standards and values of their parents on an emotional level. Such children do not find their parents' standards and values emotionally acceptable. They follow the standards to whatever extent they do simply because of pressure, not because of actual

adoption of the ideas involved. Zucker believed that the extent of adoption of parental ideals and affectional identification go hand in hand. That is, when the children really liked their parents, they adopted their ideals.

One of the first propositions to establish or demolish is that between delinquent children and their parents there is a lesser degree of affection than between nondelinquents and their parents. After establishing this, it would then be necessary to prove that such identification, or lack of it, has certain effects upon behavior that tend toward or against delinquency.

Zucker points out that we have to assume that delinquents know the difference between "right" and "wrong," hence are not in the class of "moral imbeciles." Unless this is assumed, the basis for delinquency would have to lie in intelligence. This brings up some very subtle distinctions regarding the discrepancies between "knowing" and "doing" which we need not settle at this point. It can be said, however, that in several studies no essential differences could be found between delinquents and nondelinquents with regard to knowing what constitutes acceptable social behavior.

Subjects. Twenty-five white boys who previously had been delinquent (had come into conflict with authority) constituted group A. Their mean age was 13 years and 7 months, covering a range from 11 years and 2 months to 15 years and 6 months.

The 25 subjects for the control group (group B), or nondelinquent group, had to be recruited from another school. The mean age of the group was 13 years and 8 months, ranging from 12 years and 9 months to 14 years and 9 months.

Most of the individuals in group A came from families of somewhat low socioeconomic level. In both groups most of the fathers belonged in Sims' category 4, *i.e.*, the category containing skilled laborers, members of the building trades, transportation trades, small shop owners, and personal service workers.

There was an attempt to equate the two groups (A and B) as regards intelligence. No boy having an IQ below 80 was selected. The mean IQ of group A was 92.9 and that of group B was 98.3. The range in A was from 80 to 121, and in B, from 82 to 120. In group A's school the principal, who did the selecting, was asked

to pick out only the more serious cases, whereas the principal in group B's school was asked to pick out only individuals who had not been known to commit any delinquency, and who, as well as could be determined, were normal and well-adjusted.

Procedures. The first of the several procedures involved the use of a story-completion test. Part of each of four stories were read to the subjects. The first story went something like this: Nearly two years ago, the newspapers advertised that a rich man was going to stage a foot race and give a prize to the winner. This would be a free round-the-world trip lasting about a year. The winner's expenses at the best hotels, rides in speedboats, etc., would be paid. Furthermore, the winner would be allowed to take two friends of his own age along with him. Billy entered the race and won. But then he realized he had a problem to face. It was certain that if he were to go around the world, he would have to leave his parents behind and be away from them a whole year. He wanted very much to go, but he had to think the matter over, and, ...

The third story was about like this: Jimmy went up the front steps at home without a word. There was an officer right behind him. They rang the bell, and when Jimmy's father came to the door he was shocked to see Jimmy and a policeman. The officer said, "Mr. Brown, I found your boy stealing. This is the only time you will be warned. The next time he will be taken to court." When the officer went away, Mr. Brown and Jimmy sat down and had a long talk. Jimmy was shown how wrong stealing was and what bad consequences would follow. His father pointed out that no one ever gets away with stealing and asked Jimmy never to steal again. Jimmy assured his father that he would not steal again. Two weeks later, Jimmy and a pal were walking down the street. . . .

When these stories were presented, each subject was asked to listen carefully to the instructions regarding what to do. The experimenter then said that he had four stories and was going to read only a part of each to them. The subject was supposed to finish the stories—give them an ending. The boys were allowed to respond orally or to write down the story completion.

The stories were followed by a six-section, paper-and-pencil

test given to each subject individually. In this test, additional instructions were given orally by the experimenter. Whenever any difficulty in reading the material was evidenced, help was given. The following approximates one of the six items in the first section:

	Father	*Mother*		*You*
If Jackie came home with low school grades	___	___	would yell about it.	___
	___	___	would try to make him feel better.	___
	___	___	would give him a beating.	___
	___	___	would have a talk with him.	___

The last column was to indicate what the subject would do in each case if he were the father or mother. Such a test was intended to disclose parental attitudes from the child's standpoint. It could also be expected to reveal something regarding the child's identification with one or both parents.

The second section of the test included items such as "All the kids like Bob," and "Alex gets along swell with his mother." For each of the items the boy had three choices: (1) "That's like me"; (2) "That's a little like me"; and (3) "That's not like me." The instructions were to put a cross (x) under the item that fit best in each case. The test was intended as a chance to find out how much the subject was like other fellows or was different from them.

In the third section the child was given a hypothetical situation about which to make an answer. The statement ran something like this: Let's suppose you were going to stay the rest of your life on a faraway island, which two people would you want most to take along with you?

The fourth section consisted in a table of items to check. The first column of the table was headed "Father wants me to be"; the second column pertained in the same way to the mother. The list of items in the third column were to be checked in columns 1, 2, or 4. Column 4 was headed, "I want to be." Some of the items were: polite, truthful, a fighter, quiet, generous, happy, flashy, disobedient, hot-headed, "one of the boys."

In the fifth section, 14 hypothetical results were specified as possibilities if an old Chinese magician were to give the subject

power to do anything he wanted to do. Among the suggested choices were: Get a new girl friend, change your name, have father go away on business for a few months, become a baby all over again, go out with father more often.

In the sixth and final section, a list of persons was given. The subject was to place the figure 1 in front of the person to whom it would be hardest to lie; 2 for the person next hardest, etc. The list included principal, friend, father, mother, judge, teacher, and doctor. The section also included the question: "Of everybody you know, which two of them do you admire most?" This question was thought likely to reveal the subject's identification. The answer to the last question was expected to disclose the person to whom the child had the greatest "affect" in reference to a moral situation.

Results. In the story completion, 3 of the delinquents chose to stay at home and 22 went on the world trip. Of the nondelinquents, the choices were: 18 for staying at home and 11 for taking the trip. The results of all the story completions are shown in Table 21.

TABLE 21

Completions	Delinquents	Nondelinquents	
Took trip	22	7	Story 1
Stayed at home	3	18	
Kept knives	17	6	Story 2
Did not keep knives	8	19	
Stole	19	6	Story 3
Did not steal	6	19	
Went to parent	8	20	Story 4
Went to friend	17	5	

Zucker showed that neither age nor IQ of the responding individuals had anything to do with their choices as, for example, to stay at home or go on the world trip.

In the first section of the paper-and-pencil test results, it was found that the delinquent subjects tended to identify less with their parents than the nondelinquent students did. There was no statistically reliable difference in the preference for either parent. When the delinquents and the nondelinquents were

combined into a single group, it was found that this over-all group tended to identify more often with the father.

The fathers were pictured more often by the delinquents than by nondelinquents as acting unfavorably to good relationships, in the six parts of the test. The difference was definitely reliable statistically. The differences between the mothers as described by the two classes of subjects were also reliable.

Curiously enough, both groups, the delinquents and the nondelinquents consistently stated that the mother more often reacted in a manner disturbing affectional association than did the father.

Both delinquents and nondelinquents indicated the same choices of personality traits wished for them by their parents.

Reliably more delinquents than nondelinquents did not check traits which they thought their parents esteemed.

Retaliation was indicated more often by delinquents when they felt unsatisfactorily treated by parents. This retaliation consisted in behavior of which they knew their parents disapproved.

QUESTIONS

1. Discuss distinctions and similarities in what may be meant by the terms clinical psychology and abnormal psychology.
2. State what certain authorities have believed the basis for delinquency to be.
3. What was the hypothesis that underlay Zucker's study?
4. Do delinquents know the difference between right and wrong?
5. Describe briefly what Zucker did in his investigation.
6. What is meant by identification?
7. What difference in identification was found in delinquents and in nondelinquents?
8. Did Zucker obtain reliable results from any of his tests? If so, what were they?
9. What differences did the subjects point out between fathers and mothers?
10. Point out all of the differences Zucker found between delinquents and nondelinquents themselves.

CHAPTER 47

A CLINICAL STUDY OF THINKING

Those interested in abnormal psychology have long been occupied with the problem of classification and diagnosis of individuals who deviate from normal behavior. Considerable insight into the various deviations can be gained from examining the way individuals in the various categories think. In fact the very construction of our present classification system for mental patients has involved at least a superficial recognition of differences in the way individuals think. Descriptions of the various categories in this system include statements regarding thinking, memory, perception, etc. Notwithstanding the long history of recognition of disorders in thinking, considerable information is yet to be discovered in this respect.

The work that we wish to deal with in this chapter is that of Cameron, who made a comparison of the thinking of normal adults, normal children, schizophrenes, and seniles. The study was made of thinking on the level of verbal logic and consisted in the patients completing four sentences which were presented orally to them individually. The verbatim answers were taken down in shorthand. The subjects were 22 patients with senile psychoses, 29 normal children between the ages of 7 years 2 months, and 11 years 5 months, 20 normal adults, and 25 schizophrenic patients.

Both the seniles and the schizophrenes represented selected groups, since they had to be inmates able to cooperate with the experimenter. The seniles all showed the kind of gross deficiency characteristically associated with senile deterioration. For example, only three could give their age in approximate years. The others made errors ranging from 5 to 50 years. Most of them could not say how long they had been in the hospital. Fourteen of them did not know it was a hospital they

were in. Some of them could not remember that they were married, and others gave their children's ages as greater than their own. The children used were hospitalized patients, but considered quite normally alert and cooperative. They were patients in a general pediatrics division. Defectives, neurological cases, and cases with febrile, infectious, or toxic states were excluded.

Method and Procedure. Fifteen sentence fragments were presented to the patient orally. These called for completion in the form of a causal sequence. The following are four examples:

1. I am in the hospital because . . . .
2. I am a man (woman) (boy) (girl) because . . . .
3. A man fell down in the road because . . . .
4. A fish can live in the water because . . . .

The four sentence fragments selected represent somewhat of a diversity of propositions. The first one calls for some explanation of what, to the average person, would be an indisputable fact having a fairly easy explanation. The second sentence fragment likewise deals with another indisputable fact, but allows for greater individual variation in handling. The third involves a simple hypothetical situation well within the range of everyday experience, but for which choice of causal factors may be utilized by the subject. The fourth sentence fragment calls for an explanation of a natural phenomenon usually taken for granted without explanation. The first two sentence fragments may be classified as personal, the third as other-personal, and the last decidedly impersonal.

Results. The results of this investigation can best be given in the form of quoted protocols or answers. The first answers are those of the children to the four sentence fragments.

Children's Responses.[1]

1. I am in the hospital because . . . .
   "Because I was sick."
   "Because of my eyes."
   "Because I have disease of the heart."
   "Because I got cholera."

[1] All of the protocols of the various types of subjects are quotations from Tables I to XVI of Cameron's article (A study of thinking in senile deterioration and schizophrenic disorganization. *American Journal of Psychology,* 1938 **51,** 650–664), and are used by permission of the author and the journal.

In virtually every case the children referred to their illness, and thus the answers were those that might well be expected, even from intelligent adults.

2. I am a boy (girl) because . . . .
   "Because God made me."
   "I don't know."
   "I was born that way."
   "Because I'm young."

These answers are representative of most of those given, although there were others such as, "Because I wear pants," etc.

3. A man fell down in the road because . . . .
   "It was a brick, or rather a rock; he stumbled over it."
   "Because he was riding on a horse."
   "Because he stumbled."
   "He was hit."
   "Because he was drunk."
   "Because he was weak."

A much greater variety of answers or explanations resulted from the presentation of this sentence fragment. They have to do with a variety of causes ranging from the man's not looking where he was going, the street being icy, the man being ill, his being hit by a car, to his being drunk.

4. A fish can live in the water because . . . .
   "Because it has fins or a tail."
   "Because he drinks the water."
   "Because it breathes the water."
   "It will die on land."
   "Because it was made to live in the water."
   "Because we can't live in it and he can."
   "Because it stands the water and we don't."

The completion of this sentence fragment called forth a great variety of statements, but despite their variety, they represented only a few kinds of explanation.

Normal Adult Responses. The normal adults were stenographers, attendants, and nurses (without college training) who were employed at the hospital. There was an attempt to choose those individuals whose educational and social background were roughly the same as those of the senile and schizophrenic patients.

1. I am in the hospital because . . . .
   "Because I am employed here."
   "Because I like nursing."
   "Because I could not get another job."
   "To help patients."

The reasons given by the normal adults mostly involved, as an explanation, their employment in the hospital, their interest in patients, the inability to find employment elsewhere, and their liking their employment, or advantages in medical treatment.

2. I am a man (woman) because . . . .
   "Because of the physical difference; I am a girl because I wear a dress."
   "Because I am not a man."
   "Because I was born that way."
   "Because God created me so."

Almost the same general type and variety of answers to this fragment were given by these adults as by the children.

3. A man fell down in the road because . . . .
   "Because he stumbled."
   "Because he could not stand up."
   "Because he turned his ankle."
   "Because he was careless."

According to the protocols quoted by Cameron, most of the adults suggested stumbling, slipping, or stubbing the toe. If these answers are taken as exemplary, it might be said that the adults suggested a narrower variety of explanations than the children did.

4. A fish can live in the water because . . . .
   "They have the power to breathe while under water."
   "That is its environment."
   "Because he has no lungs."
   "It's the way he is constructed. He has facilities for breathing that we don't have."

The various answers given by the adults generally concern the anatomy and physiology of the fish. A number of the verbalizations, however, were rather crude.

Senile Responses. In presenting the responses to sentence fragments made by the seniles and schizophrenes, Cameron divides them into two classes, A and B. Responses of the A type are comparable with those made by the normals. The responses of the B type indicate the pathological results. In presenting the responses of this type, an arrangement in descend-

ing order was made, the more definite pathological explanations appeared toward the bottom.

1. I am in the hospital because . . . .
   A. "Because I am sick."
      "Because my brother thought it was the place for me—if it *is* a hospital. It is better to call it a home."
      "Because I wasn't able to work."
   B. "Because I fell down just before I came here."
      "Because I couldn't wear the glasses no longer."
      "Because I'm not able, or sick, or something like that."
      "I just stepped in this morning to see if Mrs. Quiggly wants me."
2. I am a man (woman) because . . . .
   A. "Because I was made one."
      "Because the Lord made me that way."
      "Because I am a female."
   B. "Born, I guess. I don't know."
      "Because the Lord wanted me to be a woman. Of course I always have a good deal of work to do. My husband can do anything he wants to."
      "Because I am an old woman, and I have been born a good while."
3. A man fell down in the road because . . . .
   A. "Because he tripped."
      "Because he was drunk, I guess."
      "Because he got hurt."
   B. "Because he wanted to, or else he was drunk."
      "I fell down when I stumbled. I stumbled over my toe when I was nine or ten."
      "Because I think he might steal or he might be good."
4. A fish can live in the water because . . . .
   A. "Because he can't live out of it."
      "Because the water's nice for him. When he gets out of the water he dies."
      "Because it swims."
   B. "Because of being a natural water-bug."
      "Because I want to eat them and get them out." (How can they live in the water?) "They can live in the water because they are always in the water. Oh, I'm a pretty good scholar."
      "A fish has to swim. He has to get the fish, wash the fish. Everything is the main thing on Sunday there."

Schizophrenic Responses.

1. I am in the hospital because . . . .
   A. "Because I'm sick."
      "Because I had a nervous breakdown on the 28th of February, 1934."
   B. "Because I wasn't working. I'll be 28 in October."
      "I was entered in here three different times under the same name."
      "I work here for the United States Government."

"Because I love life."
"Because of an investigation of a charge and understanding of it as a condition according to Almighty God on account of a child that could not have anything on account of a constant change and abuse."
"Because I was hatched." (Why hatched?) "To separate chains from it; the papers will be there."

2. I am a man (woman) because . . . .
   A. "Because I'm born so."
   "Because I am a woman."
   B. "I am a girl because I am a she. It isn't a he or me."
   "The way I feel I am not very capable."
   "Because through some sort of a gas cubic disease, forms a quinsy disease."
3. A man fell down in the road because . . . .
   A. "Because he stumbled."
   "Because he met with an accident."
   "He may have been knocked down."
   B. "Because he was injured and shot."
   "Because he probably got electrocuted and shot and was probably dead.
   "Because of a magnet in Holland."
   "Because he might have full conscience depressed if overcrowded area; kidney disease of some kind."
4. A fish can live in the water because . . . .
   A. "Because they can't live on dry land."
   "Because it is made that way and there is oxygen in the water."
   B. "Because of sea-food."
   "Because they want to get under-seas to do what other things do.
   "Because it has learned to swim." (What if it could not swim?)
   "Not naturally, he couldn't. Why do certain gods have effects on seas like that? What does the earth have such an effect to break their backs? The fishes near home come to the surface and break." (Why?) "I think it is due to bodies that people lose. A body becomes adapted to the air. Think thoughts and break the fishes."

Discussion. The examples of the completion of the sentence fragments made by the four groups serve well to indicate what they have in common and the divergences between them. The answers given by the normal children and normal adults serve mainly as controls. Our chief interest lies in the comparisons between the seniles and the schizophrenes. The protocol material, in the B-type answers, brings out a definite set of contrasts between them.

In the answers to the first sentence fragment (regarding presence in the hospital), the seniles present plausible, even if incorrect, explanations. A kind of inventiveness, though not tallying with fact, was manifested. The schizophrenes responded with

more or less illusional material and with distortions not appearing in the protocols of the seniles.

In the completions of the second sentence fragment, the seniles again were superior. Cameron ascribes part of this to the tendency among seniles to use deistic forms of explanation quite common during the period of representing the younger days of the subjects. The schizophrenes' answers, in contrast, contained many transformations and neologisms.[1]

In explaining why the man fell down in the road, one senile out of every two gave plausible interpretations, such as that of alcoholism, injury, stumbling, weakness, or illness. Others included accounts of personal events. This tendency among seniles to intrude with information about themselves has been reported by certain other writers as characteristic. The responses of the schizophrenes are in direct contrast to this. Both the logic and the form of the responses are quite different. Among other characteristics was the tendency to include several items when one alone would suffice. For example, the schizophrenes said the man was "injured and shot," or "stumbled and was careless."

In explaining the ability of fish to live in the water, the seniles made certain misstatements, but many could be ascribed to ignorance. Others were simply irrelevant. The schizophrenes again went into a kind of loose-cluster-form expression, quite meaningless to the listener. Cameron's discussion of the two groups goes beyond the immediate findings to include certain generalities of observation regarding the two groups. These generalities do, however, tend to orient the reader toward the findings of his investigation. Cameron points out that the seniles generally (despite their enormous memory deficits and general disorientation with reference to the present) are much superior in using the ordinary facilities of communication. They maintain a type of social contact. The schizophrenes differ by being aware of who the experimenter is, where they themselves are, and

[1] Neologisms are words used with meanings that are new or not yet sanctioned by general use. Such words are, in a sense, a kind of invention, a kind of slang. Since they may have a highly restricted personal meaning, they tend to be meaningless to other people. Such words vary all the way from representing unique, quite creative inventions, to being mere gibberish.

when and where they were born, but cannot use the social tool of language nearly so well. Seniles may also express other features of socialization that are lost by the schizophrenes. The seniles may observe the simpler social amenities such as courtesy, modesty, and deference. These traits may be put to some use by senile patients who are so far lost as to be living in a bygone day, waiting for long-dead parents to come for them to take them to a home that has long ceased to exist.

Seniles retain some ability for self-criticism, and this, of course, comes to light through verbal expression. Their own deficits of uncertainty were expressed in the task of causally completing the sentence fragments. In connection with their own realization of shortcoming, they also tend to be evasive. Schizophrenes seem to miss certain points rather than to evade them.

Summary. Cameron compared statements made by selected groups of seniles and schizophrenes with those of normal children and normal adults. Neither the seniles nor the schizophrenes were able to do so well as the normals. A decided distinction between the two pathological groups was brought out in the responses. Seniles only infrequently manifested either the loose-cluster-form organization (asyndesia) or the substitution of inappropriate approximation for exact terms so typically found among schizophrenes.

QUESTIONS

1. What was the objective of Cameron's study?
2. What kinds of subjects did he use, and where were they obtained?
3. What did Cameron use as test material?
4. Was there any essential difference between the responses of children and of normal adults?
5. Identify the characteristics of the responses of the senile patients.
6. Describe the responses of the schizophrenes.
7. What are neologisms?
8. Point out the outstanding differences between senile and normal adult responses.
9. Distinguish between senile and schizophrenic "thinking."
10. Suggest additional procedures to penetrate further the differences between the thinking of the several groups used in this study.

PART XIII

INDUSTRIAL PSYCHOLOGY

been made, even without too many psychologists participating. Such events as these should be significant to the insightful in psychology and form a stimulus to further psychophysiological research and application.

The problem which we have chosen for the present chapter happens not to bear upon this specific subject, but to pertain to accidents and accident-proneness. Accidents are of many kinds and occur both in industry and out. One of the most frequent kinds of accidents is the traffic accident. Whereas we generally think of traffic accidents as not pertaining to industry, they are an important consideration in that sphere, for industry employs many drivers of vehicles.

The principles pertaining to the selection of drivers, and the question of accident-proneness as a possible human trait, can best be considered as a large problem including both industrial and nonindustrial personnel.

Industrial psychology possesses a further stake in the matter of accidents and accident-proneness by reason of the usual combining of the psychology of *business* operations with those of factory production (manufacturing). With the inclusion of business, the stake of insurance companies in understanding accident-proneness comes to be included.

In concerning ourselves with accident-proneness, we can do no better than to look to a recent review on the matter made by Johnson. While in the usual sense of the word, the material he presents is not in the form of a planned experiment of his own, he does deal with problems, hypotheses, and the available evidence bearing upon them in such ways as to clarify the subject. He was able to arrive at certain of his conclusions through the uncommon use of the tool of reasoning. Johnson's treatment of the subject shoud lead anyone to a better understanding of what to expect from various experimental techniques in attempting to select proper drivers and to screen out accident-prone individuals. Characteristically, Johnson made a good job of this task.

Procedures. Johnson points out that there are two basic procedures or methods in attempting to segregate a class of motor-vehicle drivers whose accident rate per unit time, or unit distance, is significantly greater than that of the rest of the driver

population. Procedure 1 consists in the *examination method.* In this there are two forms, the direct and the indirect. The direct form requires that each driver operate a motor vehicle in a test situation. This test would involve preferably a standard course under the observation of a trained examiner, who would rate the performance according to some scheme or other. This is a kind of "work sampling." Johnson, however, makes clear the shortcomings of this method. In fact its major defects should be quite obvious. The sampling is too small; it is obtained under conditions calling into play some nonrepresentative attitudes and motives of the driver under test. For such a method it is impracticable to set up the necessary actual emergencies testing the utmost skill of the driver, for to do so would entail the grave risk of wrecking the test car and bodily injuring or killing the occupants.

The indirect form of the examination method requires the candidate to undergo a set of performance tests pertaining to a number of supposed abilities involved in driving. These abilities pertain to attention, motor coordination and skills, traffic information, etc. Performances required in such tests are quite remote from actual driving skills and, to be of any value, must be shown by some means to be involved in driving. Not only must a general connection be involved, but scores of such tests must have a meaning in terms of driving skill and likelihood of appropriate performance under the various driving situations generally encountered on the road. The closest that the more canny investigators come to making the presumptions just stated is the assertion that their laboratory tests of attention, etc., have a high multiple correlation with accident rates. As Johnson puts it, it is preposterous to assume that a candidate's response to the laboratory situation (often taken to "symbolize" the real situation) such as a "symbolic collision," in which no fenders are dented, no blood spilt, no quarrels instigated, and no countercharges or arrests made, will indicate what the candidate will do in a real accident.

Procedure 2 is the *biographical method.* This method requires the gathering of as many data as possible regarding the personal history of each driver, including particularly those that bear upon

his accidents within a specified period. From these there is an attempt to use some constant rule whereby the data obtained give a usefully high multiple correlation with accident rate. Johnson found that, up until the time of publishing his report, this method had not been widely used, except by certain successful personnel managers who did not publish their findings. He states that ordinary workers have not yet become acquainted with how to use the technique to the best advantage.

The two general procedures just given, by their very nature, tend to supplement each other.

Johnson's next concern was an answer to the question of how good is, or what is, the predictive value of the very best test conceivable. Unless one knows the limitations of the procedures he wishes to use, he is likely to expect too much of them. Luckily it is possible to know of certain limitations beforehand. Cobb, in an analysis of this problem, gave us something useful. Cobb concluded that the number of accidents that any person has in a time period depends upon the interplay of *personal liability* with a set of events that may be labeled *luck*. The first factor, personal liability, can be broken down into contributing ingredients. Some of these could be called *intrinsic* to the personality structure of the individual. These would include physical and mental shortcomings. Others are *extrinsic* and are brought about by other persons whom he encounters. These, too, are personal in the same way that the various demands of a given vocation are personal. The luck component is independent of the individual, just as are the cards dealt to the players in a card game. If this second component did not enter into the outcome, the best-of-all tests would predict accident rate perfectly. Since chance is always an ingredient, *the ideal test is limited to sorting the individuals into liability classes, in such a way that the liability is equal for all individuals within any given class.* The test-classes, however, differ significantly from one another.

Under circumstances of this sort, the variance of the individual number of accidents from M (the mean number in each liability-class) would be equal to M itself. Likewise, the variance within the over-all distribution from its mean M would be equal to M.

We can let the total variance from the mean distribution be σ^2. The variance can be divided into two portions, the component represented by "chance" C, and $\sigma^2 - C$, the resultant of all the systematic (nonchance) factors. If R is the coefficient of correlation between the scores on the ideal test (the best possible test) and accident rate, it may be expressed, $R = \sqrt{1 - C/\sigma^2}$. Hence given the distribution of accidents per operator, one can quickly determine the predictive value of the ideal test and decide whether it is worth while to construct a practicable test. For example, Cobb, in this way, showed that for the Connecticut population, which Johnson studied, the greatest possible correlation would be .46. For another class of drivers, taken selectively from the same population, Johnson found, by Cobb's criterion, that the best obtainable correlation would be $R = .84$.

The next procedure of Johnson was to point out two features common in the field of driver testing. The first was that certain well-known investigators have claimed special usefulness for their tests or procedures in detection of good or poor drivers. Only in a few cases has statistical evidence been submitted to support these claims. The second fact true of the studies in this area, according to Johnson, was that, with certain exceptions, authors have not presented complete descriptions of large driver samples or listed a set of fundamental driver subclasses from which distributions could be reconstructed.

Johnson says that until 1941 the biographical method of finding out about the qualifications of automobile drivers was not adequately applied. Licensing authorities collected little material and did not study what they did collect. Large-scale private employers have collected data, but have either not studied it or have not published what they have found out. Johnson believes, however, that enough is known to justify the attempt to determine the limitations of the biographical method. For a certain actual occupation quite similar to driver vocations, it was established that the individual's religion, type and extent of his education, his chief interests in school, his previous jobs, his age, and the reasons he gives for wanting to train for this new occupation, etc., taken together are related to the possibility of success in training for the occupation. Johnson says that if someone pro-

duces an adequate method for combining such data some very surprisingly useful results may eventuate.

The currently common items of biographical information receiving public attention are (1) the candidate's age, (2) his own estimate of annual mileage driven, and (3) his accident rate for a given period. Item 1 can be determined. Item 2 is not one of certainty, and whether item 3 is usable depends largely upon the way accident histories are recorded. Nonfatal accidents do not seem to be always recorded.

In addition to using accident histories of one period to indicate possible accident rates in another, information about driver *ages* has been used. The grouping of ages has been various. Until recently the National Safety Council procedure of grouping all individuals younger than 21 together, and the older drivers in 10-year age classes, has been widely used. During the period of 1932 to 1936, the fatal accidents per driver were quite unevenly distributed among age classes of drivers. The rate among drivers from 16 to 20 years was 1.73 times that for the whole population. It was 1.93 times the rate of all older drivers and 2.83 times the rate of drivers between the ages of 51 and 55 years. The highest of all accident rates were manifested in the age group of 19 to 21 years. It was 3.1 times that of the 51–55-year group. The likelihood of the difference between these two groups being due to chance is roughly 10^{-1400}.

Biographical information, except for age, mileage rate, and sex of driver, has not in general been used in studies appearing in the literature. Cobb, however, showed that a battery of nine merely pencil-and-paper tests, mostly involving biographical data, gave a multiple correlation with accident rate of $R = .31$ as compared to $R = .35$ when 22 tests, some of which employed elaborate apparatus, were used. Hence the nine biographical tests eliminated more than three-quarters of the variance removed by the whole group (22 tests) and about 44 per cent of what could be expected to be removed by the ideal test. It may be deduced from this that biographical information is more practicable than the extent of its use thus far would suggest.

Johnson finally considered the matter of detection of poor and good drivers by direct and indirect examination. In the first

part of this chapter, the fact that direct examination constitutes a sampling method was stated. Johnson considered the question of whether or not an adequate sampling by direct examination could be made. By a fairly good line of reasoning, he estimated that a driver may easily encounter 1,000 accidents or near-accidents per hour on a busy highway. With 3 hrs. of driving, 3,000 such opportunities would be faced. If one drives 333.3 days per year, the total would be 1 million opportunities. If driver A is a poor one, he may have an accident one time in 100,000 chances, or 10 per year. Another driver, B, may have an accident one time in 10 million chances, or an accident or near-accident once in 10 years. Thus in the long run driver A should have 100 times as many accidents as B. A 15-min. test would allow an operator only 250 chances to have an accident. Actually, he needs 400 times this much opportunity, for he has only one accident in 10,000 on the road. The probability must also be recognized that, in the test or sampling period, the rate of accidents or near-accidents is a chance-deviation from zero of nearly 0.50. According to this line of reasoning, the sampling period is inadequate both for the good and for the bad driver.

Four important hypotheses regarding indirect examinations were tested by procedures used by Lauer, in one case, and by DeSilva and Forbes in another. These were about as follows:

1. Indirect tests may enable the users to select classes of drivers, in such a way that each class will contain enough drivers to make classification useful, and that the accident rates will differ markedly and reliably between classes. This hypothesis was verified.

2. That indirect tests may enable the user to select *individuals* whose accident rates differ widely and reliably. This hypothesis was invalidated.

3. That over an extended period, governmental authorities might advisably treat drivers selectively on the basis of their test scores. The falsification of hypothesis 2 makes this one untenable also.

4. That individual deficiencies brought to light by test results are related to types of accidents to which drivers are particularly susceptible. There are various susceptibilities such as glare. It

would be reasoned that for persons of this kind, night accidents would be more likely than for the nonsusceptible individuals.

Out of eight tests which the authors used in their studies, only one showed any association with the trait predicted. Even there the correlation coefficient, although in the right direction, was too small to be significant. Some of the test results ran counter to the predictions. Of course, counterresults could be used were they consistent, reliable, and significant.

Johnson points out the connection between this sort of indirect examining and the examining done by psychiatrists, psychoanalysts, and others whose routing practice certainly does not involve the use of statistical procedures. The deduction here is that the nonstatistical procedures would be less trustworthy than the statistical.

In conclusion, it may be said again that the intent of this chapter is to contribute something to the problem of testing for specific human traits and predicting human behavior. The student should be particularly aware of the limitations of certain assumptions and procedures, as well as the possibilities of others.

QUESTIONS

1. List as many phases of industrial psychology as you can.
2. What does the study of accident-proneness on the highways have to do with industry?
3. What is the examination method of studying accident-proneness?
4. What is the biographical method of studying accident-proneness?
5. What gross limitations in the examination method did Johnson point out?
6. What contributions did Cobb make to the study of accident-proneness?
7. What is to be expected of an ideal test to predict accident-proneness?
8. What hypotheses were made by Lauer, and by DeSilva and Forbes?
9. Which ones were verified?
10. What did Johnson have to say about indirect tests?

REFERENCES

(Most of these articles and books have been used as a basis for the foregoing chapters.)

ALEXANDER, S. J., *et al.*, 1945. Wesleyan University studies of motion sickness. *J. Psychol.* [Chap. 23]

I. The effects of variation of time intervals between accelerations upon sickness rates. **19**, 49–62.

II. A second approach to the problem of the effects of variation of time intervals between accelerations upon sickness rates. **19**, 63–68.

III. The effects of various accelerations upon sickness rates. **20**, 3–8.

IV. The effects of waves containing two acceleration levels upon sickness. **20**, 9–18.

V. Incidence of sickness at various hours of the day. **20**, 19–24.

VI. Prediction of sickness on a vertical accelerator by means of a motion-sickness history questionnaire. **20**, 25–30.

VII. The effects of sickness upon performance. **20**, 31–39.

AMES, A., JR., 1946. Binocular vision as affected by relations between uniocular stimulus patterns in commonplace environments. *Amer. J. Psychol.*, **59**, 333–357. [Chap. 12]

ANDREWS, T. G. (Ed.), 1948. *Methods of Psychology.* New York: Wiley. [Chaps. 5, 6, 7]

BACH, GEORGE R., 1945. Young children's play fantasies. *Psychol. Monogr.*, **59**, 41–69. [Chap. 45]

BARTLEY, S. HOWARD, 1936. The relation of retinal illumination to the experience of movement. *J. exp. Psychol.*, **19**, 475–485. [Chap. 19]

BARTLEY, S. HOWARD, 1941. *Vision: A Study of Its Basis.* New York: Van Nostrand. [Chap. 28]

BARTLEY, S. HOWARD, 1942. The features of the optic-nerve discharge underlying recurrent vision. *J. exp. Psychol.*, **30**, 125–135. [Chap. 26]

BARTLEY, S. HOWARD, 1943. Some parallels between pupillary "reflexes" and brightness discrimination. *J. exp. Psychol.*, **32**, 110–122. [Chap. 25]

BARTLEY, S. HOWARD, and E. CHUTE, 1947. *Fatigue and Impairment in Man.* New York: McGraw-Hill. [Chap. 40]

BARTLEY, S. HOWARD, J. O'LEARY, and G. H. BISHOP, 1937. Differentiation by strychnine of the visual from the integrating mechanisms of optic cortex in the rabbit. *Amer. J. Physiol.*, **120**, 604–618. [Chap. 28]

BILLS, A. G., 1931. Blocking: A new principle in mental fatigue. *Amer. J. Psychol.*, **43**, 230–245. [Chap. 39]

BORING, E. G., 1929. *A History of Experimental Psychology.* New York: Appleton-Century-Crofts. [Chap. 2]

BROWN, J. F., and A. C. VOTH, 1937. The path of seen movement as a function of the vector-field. *Amer. J. Psychol.*, **49**, 543–563. [Chap. 18]

BRUNER, JEROME S., and CECILE C. GOODMAN, 1947. Value and need as organizing factors in perception. *J. abnorm. soc. Psychol.*, **42**, 33–44. [Chap. 41]

BRUNSWIK, EGON, 1940. Thing constancy as measured by correlation coefficients. *Psychol. Rev.*, **47**, 69–78. [Chap. 17]

CAMERON, NORMAN, 1938. A study of thinking in senile deterioration and schizophrenic disorganization. *Amer. J. Psychol.*, **51**, 650–664. [Chap. 47]

COBB, P. W., 1940. The limit of usefulness of accident rate as a measure of accident-proneness. *J. appl. Psychol.*, **24**, 154–159. [Chap. 48]

COWLES, JOHN T., 1937. Food tokens as incentives for learning by chimpanzees. *Comp. Psychol. Monogr.*, No. 71, **14**, 1–96. [Chap. 32]

DASHIELL, J. F., 1930. Directional orientation in maze running by the white rat. *Comp. Psychol. Monogr.*, No. 32, **7**, 1–72. [Chap. 30]

DESILVA, H. R., 1926. An experimental investigation of the determinants of apparent visual movement. *Amer. J. Psychol.*, **37**, 469–501. [Chap. 18]

GIBSON, J. J., 1929. The reproduction of visually perceived forms. *J. exp. Psychol.*, **12**, 1–39. [Chap. 37]

GRAYBIEL, A., W. A. KERR, and S. H. BARTLEY, 1948. Stimulus thresholds of the semicircular canals as a function of angular acceleration. *Amer. J. Psychol.*, **61**, 21–36. [Chap. 24]

GRIFFITH, C. R., 1920. An experimental study of dizziness. *J. exp. Psychol.*, **3**, 89–125. [Chap. 22]

HELSON, H., and S. M. KING, 1931. The Tau effect: an example of psychological relativity. *J. exp. Psychol.*, **14**, 202–217. [Chap. 20]

HIGGINSON, GLENN D., 1926. Apparent visual movement and the Gestalt. I. *J. exp. Psychol.*, **9**, 228–252. [Chap. 18]

HUDGINS, C. V., 1933. Conditioning and the voluntary control of the pupillary light reflex. *J. gen. Psychol.*, **8**, 3–51. [Chap. 34]

JOHNSON, H. M., 1940. Pre-experimental assumptions as determiners of experimental results. *Psychol. Rev.*, **47**, 338–346. [Chap. 3]

JOHNSON, H. M., 1946. The detection and treatment of accident-prone drivers. *Psychol. Bull.*, **43**, 489–532. [Chap. 48]

KAY, L. W., 1943. The relation of personal frames of reference to social judgments. *Arch. Psychol.*, No. 283, **40**, 1–53. [Chap. 43]

KELLOGG, W. N., 1931. The time of judgment in psychometric measures. *Amer. J. Psychol.*, **43**, 65–86. [Chap. 38]

KUHLMANN, F., 1906. On the analysis of memory consciousness. *Psychol. Rev.*, **13**, 316–348. [Chap. 37]

Lemmon, V. W., and S. M. Geisinger, 1936. Reaction-time to retinal stimulation under light and dark adaptation. *Amer. J. Psychol.*, **48,** 141–142. [Chap. 38]

Lindsley, Donald B., 1938. Electrical potentials of the brain in children and adults. *J. gen. Psychol.*, **19,** 285–306. [Chap. 27]

Lindsley, Donald B., 1939. A longitudinal study of the occipital alpha rhythm in normal children. *J. genet. Psychol.*, **55,** 197–213. [Chap. 27]

Locke, N. M., 1938. Some factors in size-constancy. *Amer. J. Psychol.*, **51,** 514–520. [Chap. 17]

MacLeod, R. B., 1940. Brightness-constancy in unrecognized shadows. *J. exp. Psychol.*, **27,** 1–22. [Chap. 16]

Maier, N. R. F., 1939. Qualitative differences in the problem solving of normal and partly decorticated rats. *Papers Mich. Acad. Sci., Arts, Letters*, 131–145. [Chap. 31]

Maxfield, J. P., 1931. Some physical factors affecting the illusion in sound motion pictures. *J. acoust. Soc. Amer.*, **3,** 69–80. [Chap. 11]

McGeoch, John A., 1932. The influences of degree of interpolated learning upon retroactive inhibition. *Amer. J. Psychol.*, **44,** 695–708. [Chap. 36]

Menzies, R., 1937. Conditioned vasomotor responses in human subjects. *J. Psychol.*, **4,** 75–120. [Chap. 34]

Nafe, J. P., and K. S. Wagoner, 1941. The nature of pressure adaptation. *J. gen. Psychol.*, **25,** 323–351. [Chap. 21]

Nafe, J. P., and K. S. Wagoner, 1941. The nature of sensory adaptation. *J. gen. Psychol.*, **25,** 295–321. [Chap. 21]

Ogle, K. N., 1946. Theory of the space-eikonometer. *J. opt. Soc. Amer.*, **36,** 20–32. [Chap. 13]

Perkins, F. T., 1931. A further study of configurational learning in the goldfish. *J. exp. Psychol.*, **14,** 508–538. [Chap. 33]

Perkins, F. T., 1932. Symmetry in visual recall. *Amer. J. Psychol.*, **44,** 473–490. [Chap. 37]

Perkins, F. T., and R. H. Wheeler, 1930. Configurational learning in the goldfish. *Comp. Psychol. Monogr.*, No. 1, **7,** 1–50. [Chap. 33]

Proshansky, H., and G. Murphy, 1942. The effects of reward and punishment on perception. *J. Psychol.*, **13,** 295–305. [Chap. 41]

Shaw, M. E., 1932. A comparison of individuals and small groups in the rational solution of complex problems. *Amer. J. Psychol.*, **44,** 491–504. [Chap. 42]

Shurrager, P. S., and E. A. Culler, 1939. Conditioning in the spinal dog. *J. exp. Psychol.*, **26,** 133–159. [Chap. 29]

Shurrager, P. S., and H. C. Shurrager, 1946. The rate of learning measured at a single synapse. *J. exp. Psychol.*, **36,** 347–354. [Chap. 29]

Stevens, S. S., 1935. The operational basis of psychology. *Amer. J. Psychol.*, **47,** 323–330. [Chap. 3]

Stevens, S. S., 1936. A scale for the measurement of a psychological magnitude: loudness. *Psychol. Rev.*, **43,** 405–416. [Chap. 8]

SUBJECT INDEX